

UNIVERSAL
LIBRARY

OU_156605

UNIVERSAL
LIBRARY

MECHANICAL PHYSICS

by

HERBERT DINGLE

*Professor of Natural Philosophy, Imperial College
of Science and Technology (University of London)*

THOMAS NELSON AND SONS LTD
LONDON EDINBURGH PARIS MELBOURNE
TORONTO AND NEW YORK

*First published May 1942
Reprinted 1942, 1943*

PREFACE

THIS text-book is designed to present a general view of modern physics in a form suited to the needs of students of aeronautics, though it is hoped the book will also prove helpful to readers who have any of the ordinary examinations in view. At the same time, it is essentially a text-book of physics and not of aeronautics. What has been aimed at is the exposition of physical principles in a manner that will enable their application to aeronautical and related studies to be readily understood. Hence attention has been concentrated on giving a clear account of the ideas of the subject and of the results of physical research. In short, the book aims at giving a vivid general picture of the present state of physical knowledge and a practical understanding of what it means.

Physics is now such an enormous subject that comprehensiveness can be attained only at considerable cost of much that it would be desirable to achieve. It is impossible, for instance, to make the book adequate as a guide to laboratory work, and it has been thought undesirable to try to do at all what cannot be done sufficiently well.

No previous knowledge of physics is necessary to the understanding of the book, but the general attainments of the reader are assumed to be those of the post-matriculation stage. Some knowledge of elementary mechanics is essential. Mathematics has been reduced to a minimum, but the Calculus has not been avoided when much insight into physical principles can be obtained by its use. Nevertheless, it makes only rare appearances, and the student unacquainted with this branch of mathematics may omit those sections in which it appears without serious loss.

It has been my intention to make the book suitable for any of the training schemes for supplying the technical needs of the Services. The more elementary students will therefore find some portions beyond them, and the guidance of the teacher is very desirable in enabling them to make the best use of the book. It will be found, however, that knowledge of the more difficult portions has not been presumed in what follows, and the student's own attainments may be a sufficient guide to what he may pass over with safety. These remarks apply particularly to Chapter IX., which has been included because of its importance to those who can follow it, while those who cannot may simply accept the results arrived at.

Finally, a word should be said on the order of treatment adopted. The main division is between what I have called "mechanical physics" and "sub-atomic physics." At the present time it is very convenient to distinguish those departments in which the atom or molecule can still be treated as an elementary particle from those in which the structure of the atom is fundamental. Mechanical physics and sub-atomic physics form two separate volumes in this series.

The effect of this classification on the traditional division of the subject into Heat, Light, Sound, Properties of Matter, Magnetism, and Electricity, is broadly to place Heat, Sound, and Properties of Matter under Mechanical Physics, and Light, Magnetism, and Electricity under Sub-atomic Physics; but the older division is, in any case, obsolete, since many of the most vital of modern physical studies—such as radio-activity, X-rays, photo-electricity, etc.—had no place in it. Little trace of it remains here, for in each volume the aim has been to show the connection rather than the boundaries between the various physical phenomena.

H. D.

IMPERIAL COLLEGE
June 1941

CONTENTS

PART I.—PHYSICAL MEASUREMENTS

I. FUNDAMENTAL IDEAS	11
II. THE MEASUREMENT OF TIME	21

PART II.—PROPERTIES OF MATTER

III. MATTER AND EXTERNAL FORCES	49
IV. MATTER AND INTERNAL FORCES	63
V. INTERACTION BETWEEN MATTER AND MATTER	74

PART III.—HEAT

VI. HEAT AND TEMPERATURE	97
VII. THERMAL EXPANSION	105
VIII. THE MEASUREMENT OF HEAT	129
IX. RELATIONS BETWEEN HEAT AND WORK IN A GAS	145
X. THE SECOND LAW OF THERMODYNAMICS	154
XI. CHANGE OF STATE	164
XII. THE TRANSFER OF HEAT	179

PART IV.—VIBRATIONS AND SOUND

XIII. WAVE MOTION	191
XIV. SOUND AND WAVES	209
XV. THE FREQUENCY OF SOUND WAVES	217
XVI. THE REFLECTION OF SOUND	231
XVII. PRODUCTION AND DETECTION OF SOUND	237
INDEX	245

PART I
PHYSICAL MEASUREMENTS

CHAPTER I

FUNDAMENTAL IDEAS

THE SCOPE OF PHYSICS

IT is difficult to define physics in a brief statement. It is often described as the study of the properties of matter and energy, but this definition is not above serious criticism. Perhaps the least unsatisfactory definition is that physics is the study of measurable things. Most of its departments originate in simple sensations, such as those of heat and light, but at the very beginning those sensations are represented by measurable quantities which are then studied instead. Thus in heat we substitute for our direct experiences of heat and cold the readings of thermometers ; in optics we assume our sensations of sight to be caused by a physical agency called "light," whose behaviour results in things which we can measure ; in sound we substitute for direct hearing the vibrations of material bodies which we believe to be responsible for what we hear, and it is those vibrations which thereafter engage our attention ; and so on with all other physical phenomena.

At the foundation of all physical studies lies the study of motion, for this is measurable more directly than any other of our experiences. We say a body is moving when it changes its position as time goes on, and we measure its motion by the rate at which its position changes. This involves the measurement of position—*i.e.* of space—and of time.

Measurement of Space

To measure space we choose a standard rod made of some durable material, and make two marks on it. The distance between these marks is then regarded as a unit of length or of space measurement. Thus we have a standard yard in London and a standard metre in Paris, both of which are carefully preserved and kept as free as possible from disturbances. By processes which we cannot enter into here, accurate copies of these rods can be made and can be subdivided into a large number of equal parts. In scientific work the metre is the most common standard, but as it is rather long for many purposes, the scientific unit of length is taken as one hundredth part of the metre : this is called the *centimetre*. We then define the distance between two points as the number of centimetres contained in a straight line joining the points.

Measurement of Time

The measurement of time cannot be described so briefly ; we shall consider it in more detail presently. The unit of time measurement is chosen as the *mean solar second*, and this is $\frac{1}{86400}$ of the mean solar day, *i.e.* of the average time taken by the Sun to complete one circuit of the heavens as seen from a fixed position on the Earth.

There is one important difference between the measurement of space and the measurement of time. If we wish to compare two metre rods we may lay them side by side, and see if they are equal ; but we cannot so compare two seconds or two days because we never have them both at once. We must therefore *assume* that successive days are equal in duration. There is reason to believe that we should admit very slight differences—that the day, in fact, is gradually getting longer ; but the change is so exceedingly slow that it may be ignored for practical purposes.

Measurement of Motion

The motion of a body may now be measured by the distance it covers in unit time, *i.e.* in one second. This is called its *velocity* when the direction of motion also is stated. Thus, if a body moves uniformly over 100 cm. in 5 seconds, its velocity is 20 cm. a second. The body, however, may not move uniformly. For example, it may cover 10 cm. in the first second, 15 in the next, 20 in the next, 25 in the next, and 30 in the last. In that case 20 cm. a second is its *average* velocity, the actual velocity increasing steadily. The rate at which it increases is called the *acceleration*. In this example the acceleration is constant, and equal to 5 cm. a second in each second. A body may, of course, move with a variable acceleration, but we shall not need to consider such motions.

Measurement of Force and Mass

Newton's first law of motion states that a body undisturbed by any force remains at rest or moves with constant velocity in a straight line. If the velocity of a body is not constant, a force must therefore be acting on it. The force is measured by the rate of change of the velocity, *i.e.* by the acceleration, so that if a certain force gives to a body twice the acceleration given it by another, the first force is twice as large as the second.

A given force, however, will not produce the same acceleration in all bodies. We therefore say that bodies differ in something which we call *mass*, and we measure masses by the accelerations which a given force will produce in the bodies possessing them : the larger the acceleration the smaller the mass. Thus, if a certain force produces twice the acceleration in a body, A, that it produces in a second body, B, we say that B is twice as massive as A. The mass of a given piece of matter does not change however the matter be treated, except in a manner, very rarely important in practice, which is outside the scope of this book. We shall treat the mass of a body as being absolutely constant.

The unit of mass is called the *gram*, and is one thousandth part of the mass of a particular piece of matter preserved in Paris. The gram has about the same mass as one cubic centimetre of water at a temperature of 4° C. For the same kind of substance, under the same conditions, the mass is found to be proportional to the volume, so that two grams would be the mass of two cubic centimetres of water at 4° C. That is to say, if two quantities of water, at the same temperature, occupy respectively 1 cc. and 2 cc., and if they are at rest or moving with any constant velocities, the effect of subjecting them to the same force will be to produce twice as great an acceleration in the first as in the second. This enables us to define the unit of force. It is called the *dyne*, and it is that force which would produce an acceleration of 1 cm. a second in each second in a mass of 1 gm.

The relations between mass, force, and acceleration are summed up in Newton's second law of motion, which (treating mass, as we have said, as constant) may be expressed symbolically by the equation :

$$P = mf \quad \dots \quad \dots \quad \dots \quad \dots \quad (1.1)$$

where P is the force which produces an acceleration f in a mass m . If, then, m is 1 gm. and P is one dyne, f will be 1 cm. per second per second. On account of this relation between mass and force, mass is often said to represent the *inertia* of a body, *i.e.* its tendency to resist change of motion, for clearly the larger the mass the greater must be the force needed to produce a given change of motion (acceleration).

Measurement of other Physical Quantities

The measurements of length, time, and mass are regarded as fundamental, all other measurements being expressed in terms of them. We have just seen how the measurement of force can be so expressed, and the unit of force defined in terms of the units of mass, space, and time. Other important quantities closely related to force are *momentum*, defined as

the product of mass and velocity (mv), and *kinetic energy*, defined as half the product of the mass and the square of the velocity of a body ($\frac{1}{2}mv^2$). The unit of momentum, which has no special name, and the unit of kinetic energy, known as the *erg*, are appropriately taken as the momentum and twice the kinetic energy, respectively, of a mass of 1 gm. moving with a velocity of 1 cm. a second.

A force of special importance is the force of gravitation, with which, according to Newton's theory, every piece of matter attracts every other piece of matter. If m_1 and m_2 are the masses of two bodies, their gravitational attraction, F , for one another is given by

$$F = G \frac{m_1 m_2}{r^2} \quad \dots \dots \dots \quad (1.2)$$

where r is the distance between their centres of gravity, and G is a constant (the "constant of gravitation") which is equal to 6.658×10^{-8} , when F , m , and r are measured in dynes, grams, and centimetres, respectively. We are chiefly interested in the gravitation of bodies towards the Earth, exemplified by the familiar process of bodies falling downwards. m_1 is then the mass of the Earth, m_2 that of the falling body, and r is the distance between the centre of the Earth and the body. The force F in that case is called the *weight* of the body. Since the height above the Earth's surface of most bodies of practical interest is very small compared with the radius of the Earth, we may take r as the same (namely, the radius of the Earth) for all such bodies at all points of their path. The weight, W , of the body is then given by the equation :

$$W = \left(G \frac{m_1}{r^2} \right) m_2 \quad \dots \dots \dots \quad (1.3)$$

where the quantity in brackets is constant. Taking the mass of the earth, m_1 , as 5.98×10^{27} gm. and r as 6362 kilometres (approximately the value for London), its value is found

to be about 983.7. This quantity is generally denoted by g , and we have the important equation

$$W = m_2 g \dots \dots \dots \quad (1.4)$$

giving the relation between the weight and the mass of a body.

We can see at once that g is the acceleration which the force of gravity produces in the body, for, from Newton's second law of motion, $W = m_2 f$, so that $f = \frac{Gm_1}{r^2}$. Bodies falling freely under gravity near the Earth's surface should therefore experience an acceleration of about 983.7 cm. per second per second. Direct measurements of the acceleration give a value of 981.2, which shows a satisfactory agreement when allowance is made for the centrifugal force due to the Earth's rotation.

It is to be noted, however, that g is not the same all over the Earth, because the Earth is not a perfect sphere : the lower the latitude of a region, the greater is its distance from the Earth's centre, and the acceleration of gravity varies accordingly. The values given above are for the latitude of London.

When a body, on which a force, F , is acting in a certain direction, is moved through a distance d against the action of the force, *work* is said to be done, of amount $F \times d$, and it can be proved that the unit of work can be expressed in ergs, or units of kinetic energy. For example, if a weight W ($= mg$) is lifted vertically through d cm., the work done is Wd ergs, and if the weight is then released so that it falls downwards, its kinetic energy on reaching the ground (namely, $\frac{1}{2}mv^2$, where v is its terminal velocity), will be equal to Wd , apart from a slight loss due to air resistance. The work originally expended is recovered in the form of kinetic energy. This is a simple example of a very important principle known as the *conservation of energy*, which will be dealt with in more detail later.

The erg is too small a unit for convenience in many problems, and a unit of one *joule*, which is equal to 10^7 ergs, is often substituted.

If work is performed continuously for a finite time, the rate of performance becomes a matter of interest : it is known as the *power* exerted. Unit power corresponds to work done at the rate of one erg per second : no name has been given to this unit. The power corresponding to work done at the rate of one joule per second is, however, known as one *watt*.

The British System of Units.

The units of measurement which we have defined are those of the *metric* system, or *scientific* system, or *C.G.S.* system, as it is called. ("C.G.S." stands for "centimetre, gram, second," in allusion to the three fundamental units from which all others are derived). It is the system which we shall employ in the greater part of this book. Sometimes, however—particularly in British engineering—another system is used in which the units of length and mass are the foot and pound respectively, the unit of time remaining the second. The derived units in this system can be obtained at once from the definitions, and the student will have no difficulty in understanding the terms *poundal* and *foot-poundal* for the units of force and work respectively. The foot-poundal, however, (the work done when a body is moved through one foot against a force of one poundal) is less used than a unit called the *foot-pound*, which is the work done when a mass of one pound is raised against gravity through one foot. A mass of one pound is equivalent to a weight of g poundals, from equation (1.4), so that the foot-pound is equal to g foot-poundals, *i.e.* 32.2 foot-poundals, since g is 32.2 feet per second per second.

The natural unit of power would be that corresponding to one foot-poundal or one foot-pound per second. In practice, however, the unit known as *horse-power* is employed : this is equal to 550 foot-pounds per second.

Mechanics

The relations between the quantities above defined form the subject-matter of the science of *Mechanics*, which is subdivided into *Statics* and *Dynamics*. The student is expected to be familiar with the elements of these sciences. They are fundamental to the study of physics, for it has long been the aim of physicists to express the phenomena with which they are concerned in terms of mechanical quantities. They have not always succeeded, and there are still some branches of physics which it appears impossible to reduce to mechanics. Nevertheless, even in these branches, most of the phenomena which are prominent in practical applications can be described and understood most easily in mechanical terms. It is essential, therefore, that the basic principles of mechanics shall be thoroughly understood by anyone who would understand physics.

The two volumes of which this text-book of physics is composed are distinguished by the fact that the first is concerned with phenomena which can be treated directly in mechanical terms, while in the second additional concepts are necessary before mechanical treatment can be applied. Another way of expressing the difference is that in the first volume we regard matter as composed of atoms or molecules without inquiring into the structure of those atoms or molecules. In the phenomena considered they behave as units and their structure is irrelevant. In the phenomena described in the second volume, however, the structure of atoms is of essential importance, and the relations between matter and electricity dominate the treatment.

Mechanical Physics

We conclude this introductory chapter with a summary of what we may call mechanical physics. *Mechanics* proper is concerned with the relations between force, motion, etc., as they follow from the definitions of those quantities ; *mechanical physics* is concerned with the forces which are actually found

to operate within and between material bodies. Mechanics can be studied without observing the behaviour of matter, though, it is true, the mechanical problems to which most attention has been given are those which have direct application to physics. For example, the equation $P = mf$ holds good independently of the particular character of the force P , and this is a law of mechanics ; but when we observe nature we find that a particular force, varying inversely as the square of the distance between two particles, has a special importance, and the calculation and observation of the movements of bodies acted upon by such a force is a problem of physics. The outstanding example is the force of gravitation, but we shall not here consider problems of gravitation beyond what has already been said concerning weight, because since gravitation appears to be uninfluenced by any changes—such as changes of temperature, pressure, insulation, etc.—which we can make, its study is inapplicable to the practical problems with which we are concerned.

We may divide mechanical physics into four departments : (1) the behaviour of matter under the influence of external forces ; (2) the behaviour of matter under the influence of internal forces ; (3) the effect of matter on matter ; and (4) the propagation of mechanical disturbances through matter. For practical purposes, however, this classification is hardly satisfactory, because there is a large body of phenomena—heat phenomena—which to some extent belong to all the departments, but which, nevertheless, it is undesirable to consider except as a whole. We shall accordingly take the first three departments, apart from their relation to heat, and group them under the usual but not altogether satisfactory title of “Properties of Matter.” We shall then consider the subject of heat ; and end the volume with the fourth department.

In somewhat more detail the subject-matter of the four departments is as follows. In the first we have the effects of stretching and compressing, of twisting, shearing, and bending

material bodies, and the changes of both shape and size which they undergo thereby. The second includes the forces of cohesion and surface-tension. In the third come problems of impact, diffusion, and viscosity, while the fourth comprises wave and vibrational motion in matter and the phenomena of sound.

Before proceeding to carry out the programme just outlined, we have still to describe the process of measuring time. We shall do this at some length, because it is usually not given adequate treatment in books on physics, and also because of the opportunity it will give us of introducing other matters of importance to airmen concerning the movements of the heavenly bodies.

EXERCISES

1. Explain why the acceleration of a falling body varies at different parts of the Earth's surface. Where would you expect a given piece of matter to have the greater weight—at the North Pole or at the Equator?
2. A mass of 10 gm., originally supported at a height of 15 metres, is suddenly released. Find its velocity at the moment at which it reaches the ground. What is its kinetic energy at that moment? ($g=981.2$ cm. per sec. per sec.).
3. Explain why—although the acceleration produced in a body by a given force varies inversely as the mass of the body—bodies having different masses fall under gravity with the same acceleration.

CHAPTER II

THE MEASUREMENT OF TIME

THE FUNDAMENTAL PRINCIPLE

SINCE we cannot compare two different periods of time by directly measuring one against the other, the measurement of time depends on the choice of some series of constantly recurring events, between successive members of which we can regard the periods of time as equal. Thus we may—as Galileo did on one occasion—take our pulse-beats as occurring regularly, and choose the unit of time as the interval between two successive beats. This, however, would be impracticable because the rate would vary from one person to another, and no consistent results would be obtainable. The most suitable series of events is provided by the apparent movements of the heavenly bodies, and these are chosen to furnish us with a scale of time measurement. It is necessary, therefore, to understand something of the movements of the stars before we can appreciate their use for this purpose.

The Apparent Movements of the Stars

The Earth rotates continuously on its axis, and revolves continuously round the Sun. The first movement is visible to us as a regular circulation of the heavenly bodies. They appear as though they were situated on a large sphere whose centre is at the Earth and whose axis of rotation is a prolongation of the Earth's axis. It is convenient to suppose that this sphere, which is called the *celestial sphere*, actually exists, and that the stars are fixed points on it. We can then imagine

that we are at rest, and that the celestial sphere revolves round us once a day. Fig. 1 shows the celestial sphere with the Earth at its centre. The Earth's equator cuts it in the *celestial equator*, and the Earth's axis meets it at the *north* and *south celestial poles* (*P* and *P'*).

Let us imagine ourselves watching the sky at night, with our attention fixed on a particular star which we observe to rise in the east. It crosses the sky, getting higher above the horizon until it is due south, when it begins to sink again, and finally sets somewhere in the west. If the weather is

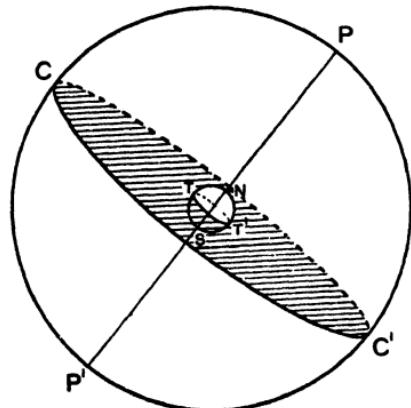


FIG. 1

The Earth and the Celestial Sphere
P, P' North and South Celestial Poles
CC' Celestial Equator
N, S North and South Terrestrial Poles
TT' Terrestrial Equator

favourable we can observe the same thing happening the next night and on succeeding nights, and we notice that the star performs a regular series of revolutions over the same path in the sky. Somewhat similar movements are shown by other stars, and since the pattern formed by the stars remains constant, their only apparent movement being this daily revolution, we are justified in imagining for practical purposes that the movement is due to the revolution of a sphere on which they are fixed.

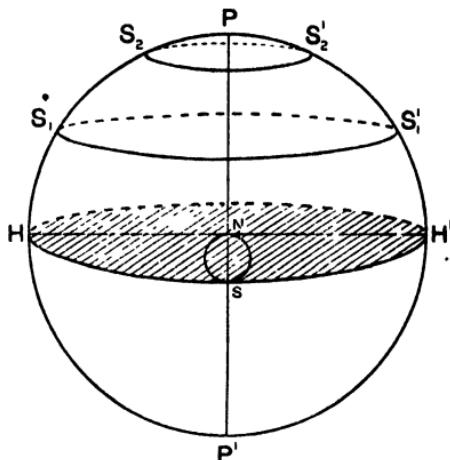
The apparent movements of the stars, however, are not the same for observers all over the Earth. To an observer at the north pole, the stars in the northern half of the sky appear to move in horizontal circles, while those in the

southern half of the sky are never seen. This is illustrated in Fig. 2, in which the plane of the horizon (the tangent plane

FIG. 2
Revolution of the Celestial Sphere as seen from the North Pole

HH' Horizon (Parallel to Equator)

$S_1 S'_1$) Apparent paths of stars
 $S_2 S'_2$)



to the Earth at the position of the observer) is shown by a shaded circle, and the paths of the stars, which are circles centred in the axis of the sphere, are clearly parallel to the

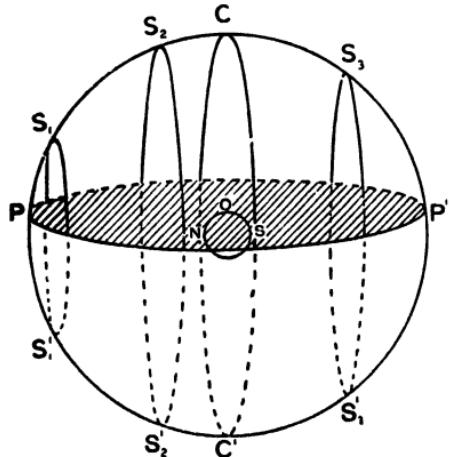


FIG. 3

Revolution of the Celestial Sphere as seen from the Equator

P, P' Celestial Poles (on horizon)

N, S Terrestrial Poles

O Position of observer

CC' Celestial Equator

$S_1 S'_1$)

$S_2 S'_2$) Apparent paths of stars
 $S_3 S'_3$)

plane of the horizon. At the Earth's equator, on the other hand, as is shown by Fig. 3, the stars appear to move in

vertical circles, all the stars in the sky being above the horizon for twelve hours out of the twenty-four.

At an intermediate place, such as London, the apparent movement of a star depends on its distance from the north celestial pole. Stars near the pole are seen to move in inclined circles (such as A_3B_3 in Fig. 4), remaining always above the horizon, while those (e.g. A_4B_4 in Fig. 4) near the south celestial pole are never seen. (In southern latitudes, of course, it is the stars near the *south* celestial pole that appear to move in inclined circles, and those near the *north* celestial pole that

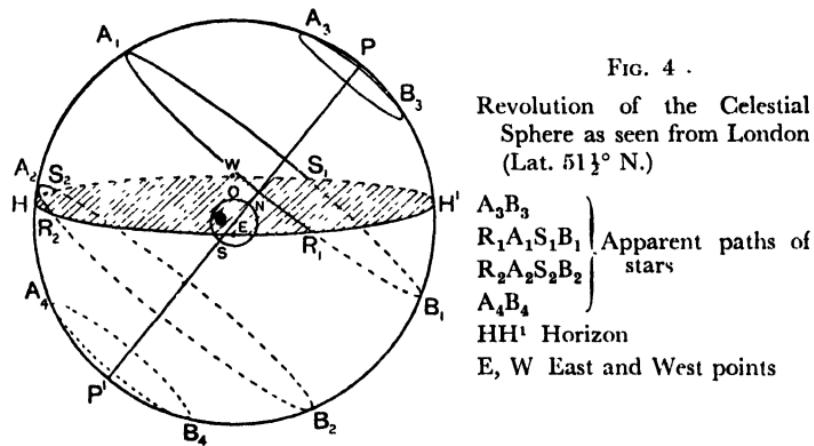


FIG. 4 .

Revolution of the Celestial Sphere as seen from London (Lat. $51\frac{1}{2}^{\circ}$ N.)

A_3B_3
 $R_1A_1S_1B_1$
 $R_2A_2S_2B_2$ } Apparent paths of stars
 A_4B_4
 HH' Horizon
 E, W East and West points

are never seen.) Such stars are said to be *circumpolar*. Other stars are above the horizon during part of their revolution and below during the other part; typical paths of such stars are $R_1A_1S_1B_1$ and $R_2A_2S_2B_2$ in Fig. 4. It is one of these stars that we pictured above as rising in the east and setting in the west. Clearly, however, the points of rising and setting are not necessarily due east and due west, for the paths shown in Fig. 4 meet the horizon in widely separated points. R_1 and R_2 are the places where the stars rise, and S_1 and S_2 the places where they set, while E and W are the east and west points.

There is one thing, however, which is common to all the

stars, no matter from what place they are observed, and that is that (apart from extremely small differences, which we shall ignore) they all take the same time to complete a revolution. If, then, we set a telescope in a certain direction and observe a star crossing its field of view, the interval which elapses before the same star appears again will be the same, whatever star we select. This interval is chosen to give us our fundamental measure of time : it is called *one sidereal day*. The twenty-fourth part of the sidereal day is the sidereal hour, and this is divided into 60 sidereal minutes, each of which is made up of 60 sidereal seconds.

The Apparent Movements of the Sun

The astronomer generally uses the sidereal time system just described, but for ordinary civil purposes it is unsuitable because our daily life depends so much on the Sun. Now the Sun, although it is itself an ordinary star, does not seem to remain fixed on the celestial sphere as the other stars do, because of the Earth's annual revolution round it. On account of this motion it appears to describe a great circle on the celestial sphere once every year. It shares, of course, in the general diurnal motion of the sphere, but if, on a particular day, it appears in a certain direction (say, due south) at the same time as some star, it will appear in that direction the next day a little later than that star (about four minutes later) because, during the revolution of the sphere it has crept eastwards a little. The next day it will be about twice as late, and the difference accumulates as the days go on until after a year the Sun has completed the whole circuit of the sphere and returned to the star in question again.

One result of this is that we do not see the same night sky in the summer as in the winter. For the stars which we see at night are, of course, those in the part of the sky where there is no Sun, and this part continually changes as the Sun makes the circuit of the heavens.

The Ecliptic : It is found by observation that the course which the Sun takes during the year is not along the equator but along a circle inclined to the equator at about $23\frac{1}{2}$ degrees. This circle is called the *Ecliptic*. We can easily see from Fig. 5 that it is the circle in which the plane of the Earth's orbit round the Sun cuts the celestial sphere, for the apparent creeping of the Sun over the sphere is simply an appearance due to the revolution of the Earth round the Sun.

The ecliptic cuts the equator in two opposite points, known respectively as the *First Point of Aries* (Υ) and the

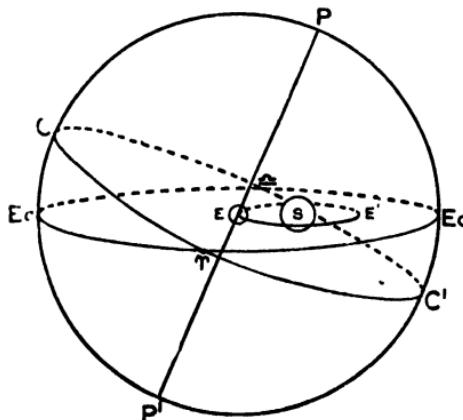


FIG. 5
The Celestial Sphere and the Ecliptic

CC' Celestial Equator

EcEc' Ecliptic

EE' Earth's Orbit

S Sun

E Earth

Υ First point of Aries

\simeq First point of Libra

First Point of Libra (\simeq). Aries and Libra are the constellations, or groups of stars, in which these points were once situated (for reasons outside our scope their positions change very slowly), and the symbols in brackets are used to represent them. A narrow zone of the sky, extending a short distance on each side of the Ecliptic, is called the *Zodiac*. It is divided into twelve equal parts, known as the Signs of the Zodiac, which are given the following names and symbols :

Aries (the Ram) . . .	Υ	Cancer (the Crab) . . .	∞
Taurus (the Bull) . . .	\wp	Leo (the Lion) . . .	\wp
Twini (the Heavenly wings) . . .	Π	Virgo (the Virgin) . . .	\wp
		Libra (the Scales) . . .	\simeq

Scorpio (the Scorpion)	11	Aquarius (the Water-Bearer)	22
Sagittarius (the Archer)	17	Pisces (the Fish)	26
Capricornus (the Goat)	19		

The first use that we make of the ecliptic is to obtain from it the starting-point for the sidereal day. We have seen that the *length* of the sidereal day is the time of revolution of any star, but the moment when the day is said to begin depends on the star chosen and the direction in which we observe it. We choose to observe due south, and we choose to mark the beginning of the day, not the moment when a star, but the moment when the First Point of Aries appears there. There is no star at this point, so we cannot directly observe it, but from observations of the stars we can calculate just when the First Point of Aries will be there, and time our clocks accordingly.

When a star, or any other point on the celestial sphere, in its daily revolution appears due south, it is said to "cross the meridian." The *meridian* is a great vertical circle in the sky, passing through the point directly over the observer's head (the *zenith*), and the celestial poles. It is, of course, a different circle for different observers because, although the celestial poles are fixed points on the sphere, the zenith depends on the observer's position on the Earth. The sidereal day, therefore, begins at different times for observers at different places. For each of them it begins when the First Point of Aries crosses the meridian, but they have different meridians.

Position on the Celestial Sphere

It is important to be able to state precisely the position of any star or other point on the celestial sphere. For this purpose we use *co-ordinates*, of which several systems have been devised. It will be sufficient for our purposes to describe three of them.

The first is almost identical with the system used on the Earth, in which the position of a point is indicated by its latitude and longitude. Imaginary great circles, called *meridians of longitude* are drawn through the Earth's poles, and circles, called *parallels of latitude*, are drawn parallel to the equator. Starting from the meridian of longitude passing through Greenwich (a meridian chosen arbitrarily, but universally adopted when once chosen), we call the *longitude* of a point the angular distance (east or west) from this meridian to the meridian of longitude which passes through the point ; and we call the *latitude* of the point the angular distance (north or south) from the equator to the parallel of latitude which passes through the point. When the longitude and latitude are stated, the position of the point on the Earth's surface is definitely fixed.

On the celestial sphere we similarly draw great circles (called *hour-circles*) through the poles, and a series of circles (called *parallels of declination*) parallel to the equator. The First Point of Aries is chosen as the " Greenwich " of the sky, and the co-ordinates corresponding to longitude and latitude on the Earth are called *Right Ascension* and *Declination* respectively. Thus, in Fig. 6, the right ascension of the Point A is γB and its declination is BA .

Declination, like latitude on the Earth, is counted in degrees, and may be north or south : thus we may have a declination of $N.60^\circ$ or $S.20^\circ$. But whereas terrestrial longitude is usually reckoned in degrees east or west, right ascension is always measured eastwards and, though occasionally stated in degrees, is usually given in hours, minutes, and seconds. The convenience of this will be seen when we come to the actual process of measuring time, but it should be understood at once that the word " hour " here is used primarily as a unit of angle only : time is not yet concerned in the matter. An hour is simply a name for an angle of 15° , hence 24 hours are equal to 360° , corresponding to the fact that the celestial sphere takes 24 sidereal hours to rotate through

360°; and a minute and second are angles which are respectively $\frac{1}{60}$ and $\frac{1}{3600}$ part of an hour. Unfortunately the words "minute" and "second" are also used to denote different measures of angle, namely, the $\frac{1}{60}$ and $\frac{1}{3600}$ part of a *degree*. This is a possible source of confusion, to minimize which the latter measures are often called "minutes of arc" and "seconds of arc." The signs ' and " are used to indicate them (e.g. $60^{\circ} 45' 30''$ stands for an angle of 60 degrees, 45 minutes of arc, 30 seconds of arc), while the signs "m." and "s." are used for minutes and seconds in

FIG. 6

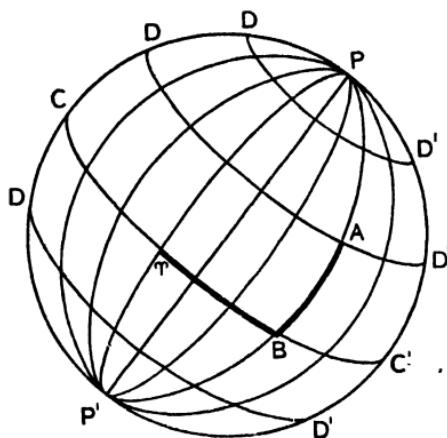
Hour circles and parallels of declination

CC' Celestial Equator

♈ First point of Aries

♉ B Right Ascension of point A

BA Declination of point A



the other sense (e.g. 12 h. 15 m. 20 s. stands for an angle of 12 hours, 15 minutes, 20 seconds, which the student can easily prove is equal to $183^{\circ} 50'$).

The second system of co-ordinates, less frequently used than that just described, is one in which the ecliptic instead of the equator is taken as the fundamental circle, and great circles through the "poles of the ecliptic" (two points each 90° from all points of the ecliptic, just as the celestial poles are each 90° from all points of the equator) are taken instead of hour circles. The angle between two such circles, one passing through ♈ and the other through any point (the angle being measured eastward from the first), is called

the *celestial longitude* of the point, and the angular distance of the point north or south of the ecliptic is called its *celestial latitude*. It is unfortunate that these names should be so similar to those of terrestrial co-ordinates which resemble them less than they do the co-ordinates of the other system, and the student must learn to avoid confusion in this matter. In Fig. 7, Q and Q' are the poles of the ecliptic, and the co-ordinates of the point A on the two systems are as follows : right ascension = πB , declination = BA ; celestial longitude = πD , celestial latitude = DA .

Since, in both these systems, the position of the point

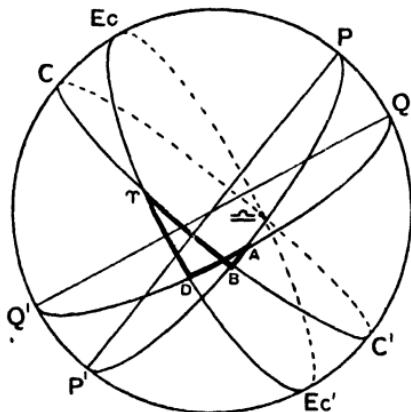


FIG. 7
Co-ordinates on the Celestial Sphere

- CC' Celestial Equator
- EcEc' Ecliptic
- PP' Celestial Poles
- QQ' Poles of the Ecliptic
- πB Right Ascension of A
- BA Declination of A
- πD Celestial Longitude of A
- DA Celestial Latitude of A

is referred to circles fixed on the celestial sphere, the co-ordinates are independent of the position of the observer on the Earth. In the third system, now to be described, the co-ordinates vary with both the place and the time of observation. The fundamental circle here is the *horizon*, *i.e.* the great circle in which the celestial sphere is intersected by a plane tangential to the Earth at the position of the observer. The variation of the horizon with the observer's position is shown by comparing it with the position of the poles in Figs. 2, 3, and 4. The radius of the celestial sphere is taken to be infinite, so that the Earth shrinks to a point, and we may therefore equally well define the horizon as the great

circle in which the celestial sphere is cut by a plane through the Earth's *centre* parallel to the tangent plane.

One "pole" of the horizon is obviously the *zenith*, the point directly overhead, and the other, which is directly underfoot, is called the *nadir*. A great circle on the celestial sphere through these points, and therefore perpendicular to the horizon, is known as a *vertical circle*. The particular vertical circle which passes through the celestial poles is, as we have said, called the *meridian* of the place of observation, and the vertical circle at right angles to this is called the *prime vertical*.

FIG. 8

Altitude and Azimuth

Z Zenith

ZENaW Prime Vertical

Na Nadir

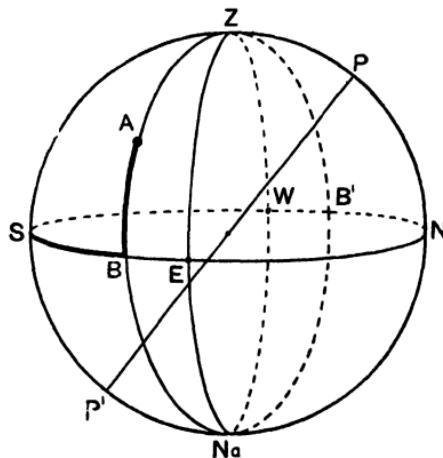
ZBNaB' Vertical Circle

N, S, E, W North, South, East,
and West points

SENW Horizon

SZNNa Meridian

{BA Altitude of A
SB Azimuth of A



The points in which the horizon is cut by the meridian are called the *north and south points*, respectively, and those in which it is cut by the prime vertical are the *east and west points* (see Fig. 8).

The co-ordinates of a point on this system are, then, the *altitude*, which is the angular distance of the point above or below the horizon, and the *azimuth*, which is the angular distance from the meridian of the vertical circle passing through the point. The azimuth may be reckoned from either the north or the south point, but the point selected must always be stated. In Fig. 8, for example, the altitude of the point A is AB, and its azimuth may be stated as any

of the four angles SB, SWNEB, NEB, NWSB. If the angle SB is 60° , for example, these four ways of describing the azimuth are written S. 60° E., S. 300° W., N. 120° E., and N. 240° W. The initial letter states the point from which, and the final letter the point towards which, the angle is measured.

Sometimes the *amplitude* is stated instead of the azimuth. This is a measurement along the horizon from the east or west instead of from the north or south point. The amplitude of A may accordingly be expressed as E. 30° S., E. 330° N., W. 150° S., W. 210° N.

It is easy to see that the altitude and azimuth of a point vary with the place of observation, because the horizon is a different circle for each position on the Earth. It is obvious also that these co-ordinates change with time, even at the same place, because the celestial sphere is continually revolving round the Earth. We can see the variation in altitude and azimuth of a star, for example, as it rises, crosses the sky, and sets.

Altitude of the Pole

There is a very important relation between the position of the observer and that of the celestial pole ; namely, that the altitude of that celestial pole which is above the horizon is equal to the observer's terrestrial latitude. Thus, at London, where the latitude is $51\frac{1}{2}^\circ$ N., the north celestial pole is $51\frac{1}{2}^\circ$ above the horizon, and, being a fixed point, remains at that altitude throughout the day. This may easily be seen from Fig. 9. Here α is the altitude of the pole, and γ is the terrestrial latitude of the observer at O. By adding the angle OC_1P to either α or γ we obtain a right angle, so that $\alpha = \gamma$.

Hour-Angle

Let us imagine ourselves standing at a fixed point on the Earth and looking towards the north or south point. Our

meridian then arches the sky through the north and south points and passes directly over our heads (through the zenith). It remains fixed with respect to us, and the celestial sphere in its revolution passes continuously across it. Consider an hour-circle passing through a particular star : this is, of course, fixed with respect to the sphere. Two points of this hour-circle, namely, the celestial poles, remain fixed with respect to us, for they are on the meridian ; and the hour-circle, with the star on it, revolves round an axis through these points. When the star is on the meridian the hour-

FIG. 9

Altitude of the Pole

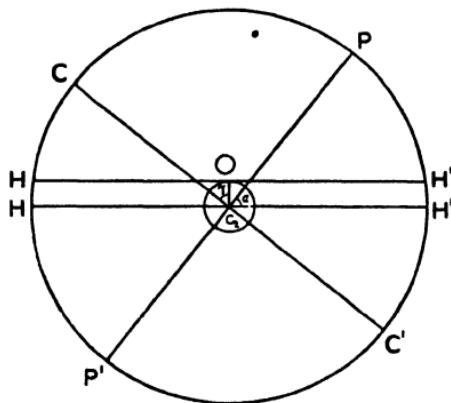
P, P' Celestial Poles

C, C₁ Earth's Centre

O Observer

CC' Plane of Terrestrial and
Celestial Equator

HH' Direction of Horizon



circle coincides with the meridian, and thereafter it makes a gradually increasing angle with the meridian. This angle is called the *hour-angle* of the star.

Sidereal Time

Consider now the First Point of Aries. When this is on the meridian its hour-angle is zero and the sidereal day begins. Thereafter the hour-angle increases until 24 sidereal hours later the First Point of Aries is back on the meridian again. We use the hour-angle of the First Point of Aries to measure hours, minutes, and seconds (in *time* measure) of the sidereal day. Thus, when the hour-angle is 15° (i.e. one hour in the other system of measuring angles) we say that one hour of

sidereal time has elapsed, and so on. We may therefore define the sidereal time of day as "the hour-angle of the First Point of Aries." In Fig. 10 the hour-angle of Υ is $MP\Upsilon$ —or, which is the same thing, the angular distance $M\Upsilon$ measured along the equator. This is the sidereal time at the instant represented. As $P\Upsilon$ sweeps round the axis the sidereal time advances steadily.

Now imagine a star A (Fig. 10). The hour-circle passing

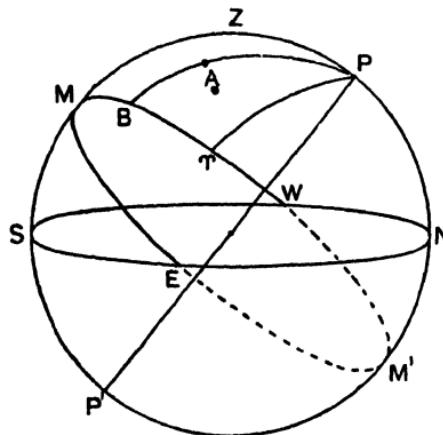


FIG. 10
Hour-Angle and Sidereal Time
SENW Horizon
EMWM' Equator
PSP'N Meridian
N $P\Upsilon$, PB Parts of Hour-circles
 $MP\Upsilon = M\Upsilon$ Sidereal Time
 $\Upsilon PB = \Upsilon B$ Right Ascension of Star A
 $MPB = MB$ Hour-angle of Star A

through it meets the equator at B. MB is then the hour-angle of the star, and $B\Upsilon$ is the right ascension of the star (cf. Fig. 6). Now obviously :

$$M\Upsilon = MB + B\Upsilon. \dots \dots \dots \quad (2.1)$$

so that

sidereal time = hour-angle of star + right ascension of star,

and this applies to any star at all. We can therefore determine the sidereal time at any instant by observing a star whose right ascension is known ; and this overcomes the difficulty arising from the fact that there is nothing visible in the sky to mark the First Point of Aries. The right ascensions of the brighter stars have been measured very accurately, and are tabulated in the *Nautical Almanac*. The process of measuring

sidereal time therefore resolves itself into that of observing the hour-angle of a star, and adding its right ascension.

For simplicity we choose the moment when the hour-angle of the star is zero ; *i.e.* when the star is on the meridian. A specially mounted telescope (known as a *transit circle*) is fixed so that it can revolve in the plane of the meridian about a horizontal axis, and can therefore be set to observe any star that crosses the meridian above the horizon. By a special device the reading of the observatory clock is observed at the instant when the star crosses the meridian. By equation (2.1), putting $MB = 0$, we see that the sidereal time at that instant is the right ascension of the star. The difference, therefore, between the reading of the clock and the tabulated right ascension of the star is the error of the clock. In observatories this observation is made at frequent intervals, and the error of the clock is recorded. No attempt is made to keep the clock exactly right. All that it is necessary to know is the exact amount by which its readings must be corrected.

Solar Time

The observations just described form the basis of our system of time measurement. For the astronomer they are sufficient : he is quite satisfied with the sidereal time system. But for ordinary life, as we have said, it is unsatisfactory because, owing to the movement of the Sun along the ecliptic, we should have the day beginning sometimes when the Sun is visible on the meridian and at other times when it is below the horizon, and much confusion would result.

The simplest modification to make is to substitute the Sun for the First Point of Aries. We might then say that the day begins when the Sun is on the meridian, and that the time of day is the hour-angle of the Sun. This would be quite satisfactory if it were not that the rate at which the right ascension of the Sun changes is irregular in sidereal time units. This rate is about 4 m. a day, but it is sometimes less and sometimes more than this amount. This would not matter

if we could easily make clocks which would record the solar time intervals. We could use such clocks for civil life and leave sidereal time clocks to the astronomers. But sidereal time seems to be, in a sense, more "regular" than solar time, and while we can make fairly simple mechanical devices for indicating sidereal time, we cannot do so for solar time. We are therefore forced to say that the rate of motion of the Sun over the celestial sphere is irregular.

The problem now to be solved is this. We must have a time system based on the movements of the Sun, but it must be one in which the unit intervals (days, hours, minutes or seconds) must all have the same length in sidereal time units. The solution is to measure the length of the solar day (the time between successive meridian passages of the Sun) by the sidereal clock for every day throughout the year, and take the average value: this is called the *mean solar day*. We then imagine a body called the *mean Sun*, which moves along the equator at the average rate of motion of the real Sun along the ecliptic, and we define *mean solar time* as the *hour angle of the mean Sun*. Mean solar time is then quite uniform with respect to sidereal time, and also the hour-angle of the actual Sun at any given mean solar time is never very different on one day from what it is on another.

There are thus two systems of solar time. In the first, which is called the *apparent solar time* system, the day begins when the actual Sun crosses the meridian, and the apparent solar time of day is the hour-angle of the actual Sun. This is never used in practice, although, since all solar observations have, of course, to be made on the actual Sun, apparent solar time frequently occurs in astronomical calculations. The other system—the *mean solar time* system—is that which we use in everyday life. The time of day on this system is the hour-angle of the mean Sun, and the day begins when the mean Sun crosses the meridian. The particular meridian passage chosen, however, is not that which occurs above the

horizon—that corresponds to noon, half-way through the day—but the passage below the horizon at midnight. Astronomers prefer to start their day at noon, but it is inconvenient in civil life for the date to change in the midst of the daily activity.

We can now see why the hour is a convenient unit of angle. By specifying right ascensions in hours we can apply equation (2.1) directly, and the hour angle of the Sun or the First Point of Aries, when so expressed, gives the corresponding time immediately.

Determination of Time

We may summarize the process of determining mean solar time as follows. A transit circle is set up, and the times shown by a sidereal clock are observed for the passage of several stars across the meridian. The differences between these clock times and the right ascensions of the corresponding stars give the error of the clock. The time of passage of the Sun across the meridian is then observed by the clock and corrected for the error, for each possible day of the year, and the mean value of the apparent solar day, in sidereal time units, is thus determined. This is the length of the mean solar day. We can then construct a mean solar clock, which is essentially the same as a sidereal clock rated to run rather more slowly. Choosing some convenient time at which the mean Sun is assumed to be coincident with the actual Sun, we can then set the mean solar clock right for that moment and thereafter determine its error from observations of the stars or of the real Sun.

The Equation of Time

The difference between the apparent solar time and the mean solar time (apparent *minus* mean) at any instant is called the *equation of time*. If we correct our mean solar clock from observations of the real Sun, we must know the equation

of time for each moment throughout the year. We cannot here enter into the way in which this is calculated, but it may be said that the reason for the irregularity of the real Sun is twofold. First, the Sun does not move uniformly along the ecliptic because the Earth does not revolve uniformly around the Sun ; it travels faster in winter, when it is nearer the Sun, than in summer, when it is farther away. Secondly, since the Sun travels along the ecliptic and not along or parallel to the equator, its hour-angle would not grow uniformly even if its motion along the ecliptic were

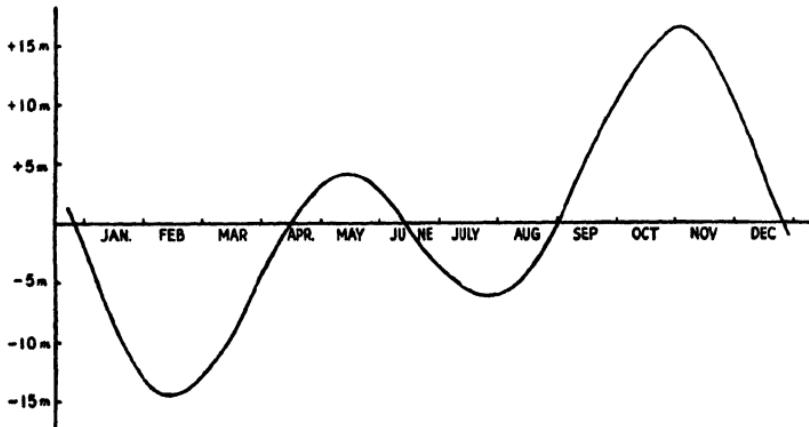


FIG. 11

The Equation of Time (apparent Solar Time — mean Solar Time)

uniform. From detailed consideration of these two effects we can calculate the equation of time and so use our observations of the actual Sun to determine mean solar time.

Fig. 11 shows how the equation of time varies throughout the year. At most it amounts to some seventeen minutes. A sundial—which, of course, records apparent solar time—may therefore give readings which differ by so much as this from our clock readings. The exact value of the equation of time is given in the *Nautical Almanac* for every day throughout the year.

Local and Standard Time

If we were confined to one spot on the Earth's surface, the mean solar time system as so far described would need no further elaboration. Consider, however, what would happen if we moved to a place of different terrestrial longitude. Since the meridian of the new place would be different from that of the old, the mean solar time (and, of course, the sidereal time and apparent solar time also) would be different, for the day would start at each place when the mean Sun came to the corresponding meridian. We would therefore have to change our watches continuously, moving them forwards as we journeyed eastwards and back as we journeyed westwards. In crossing England, for instance, the change would amount to more than a quarter of an hour. This would be very inconvenient. We therefore give up the system of *local time* (as time referred to the meridian of any given place is called), and for the whole of the United Kingdom we adopt the meridian of Greenwich. It is then unnecessary to change our watches as we move from place to place, but, to compensate for this, the clock times of rising, southing, and setting of the Sun vary as we travel eastwards or westwards.

If we go to greater distances, however, this device becomes impracticable. It is desirable to have "12 o'clock noon," for example, not far from the middle of the day, but if the meridian of Greenwich were adopted for the whole Earth, this hour would occur in the middle of the night at some places and at sunrise or sunset at others, and only places not far from the meridian of Greenwich would experience all the advantages of a system of solar time. Accordingly we compromise between a system of local time, based on any meridian we choose, and a system based on the single meridian of Greenwich, and choose certain *standard* meridians, mostly separated by intervals of one hour (15°) of longitude. Each region of the Earth, then, uses a mean solar time system based

on a standard meridian passing near it. Thus, for central Europe, the standard meridian is one hour east of the meridian of Greenwich, and when we go there we must advance our watches abruptly by one hour on crossing the boundary between the regions adopting the Greenwich and central European meridians respectively. On the other hand, when we journey westwards, our watches must be put back, and by the time we have reached the north-east coast of North America they are four hours behind Greenwich time.

The Date-line

Consider, now, two travellers who start out from Greenwich, one going east and the other west. Their watches show the same time at starting, but, as they journey on, the first puts his watch forward at successive stages, and the other similarly puts his watch back. By the time they have reached the 180th meridian from Greenwich the first has advanced his watch by twelve hours, and the second has retarded his by twelve hours, so that if they meet there they will find that one is a whole day in advance of the other. This meridian is called the *date-line*, and ships crossing it have to change their date by a whole day. A ship travelling from east to west loses a day, its time changing, say, from 6 a.m. on Monday to 6 a.m. on Tuesday in a single instant. A ship crossing the line from west to east, on the other hand, repeats a day, and goes back from 6 a.m. on Tuesday to 6 a.m. on Monday again. Fortunately this meridian passes almost entirely across the ocean, so that most of the complications which would arise if it crossed a densely populated region are not realized.

Summer Time

During summer months the clock is advanced by one hour in some countries. This is equivalent to choosing a meridian one hour east of that which fixes the standard time of the

region in question.* The change is made for economic reasons, and is effective solely because our active hours have gradually drifted to a later and later period of the twenty-four hours forming a mean solar day. Our "midday" meal, for instance, is usually taken at 1 o'clock or later, and not at noon, and we retire to rest an hour or two before midnight, rising several hours after midnight. In a life ordered according to nature, our hours of sleep would be distributed equally on either side of midnight. So tenacious are we of our habits that it is found easier to tamper with the clock than to interfere with them.

The introduction of Summer Time is sometimes thought to be an interference with nature. This is a mistake. We cannot interfere with nature in these matters ; things take their course irrespective of our whims or devices. We can adjust our arbitrary records of time to conform to our habits, or we can change our habits in order to maintain our arbitrary system of time-recording. That is our own affair. We neither gain nor lose a single moment of sunshine by such practices, just as we neither gain nor lose a day when we change the *name* of the day on crossing the date-line.

The Seasons

Our knowledge of the celestial sphere now enables us to understand the phenomena of the seasons—in particular the variation of the hours of daylight throughout the year. We may consider this from two points of view. First, we may think only in terms of the apparent movement of the Sun on the celestial sphere ; and, secondly, we may think of the revolution of the Earth about the Sun, which is responsible for that apparent movement.

In Fig. 12 the equator and ecliptic are shown, as well

* It has recently (1941) been decided, as an emergency measure, to adopt in Great Britain a meridian two hours east of the Greenwich meridian during the central part of the summer. This new time system has been called "Double Summer Time."

as two parallels of declination passing through the points of the ecliptic which have the greatest northerly and southerly declinations. These points are respectively $23\frac{1}{2}^{\circ}$ north and south of the equator. It is clear that the Sun changes its declination during the year from $23\frac{1}{2}^{\circ}$ N. to $23\frac{1}{2}^{\circ}$ S., and back again.

Let us now consider how the declination of a body affects its positions of rising and setting. In Fig. 13, CC' represents the equator, and it is clear that all points on the equator rise and set at the east and west points respectively. AA' represents the daily path of a point whose declination is northerly.

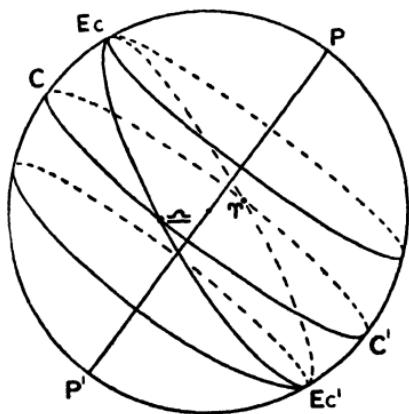


FIG. 12

Equator and Ecliptic
 P, P' Celestial Poles
 CC' Celestial Equator
 EcEc' Ecliptic

The points of rising and setting are R₁ and S₁, and it will be seen that the greater part of the path of the point is above the horizon. If the Sun has a northerly declination, then, the day will be longer than the night. On the other hand, we can see from BB', which represents

the path of a point whose declination is southerly, that such a point is below the horizon longer than it is above, and if the Sun is there the day will be shorter than the night.

The foregoing remarks apply to observations in northern terrestrial latitudes, where the north celestial pole is above the horizon. In southern latitudes the points of southerly declination will spend most of their diurnal circulation above the horizon. Hence, in northern latitudes we have long days and short nights when the Sun's declination is northerly, and short days and long nights when its declination is southerly.

In southern latitudes this is reversed, so that the northern summer is the southern winter, and *vice versa*. The greater the declination of the Sun, the more pronounced is the effect, so that when the Sun has a declination of N. $23\frac{1}{2}^{\circ}$ and S. $23\frac{1}{2}^{\circ}$, we in northern latitudes have our midsummer and midwinter days respectively. The points on the ecliptic at which the Sun is then situated are called the *solstices*. The summer solstice occurs about 22nd June and the winter solstice about 22nd December.

When the Sun is on the equator, the day and night, as

FIG. 13

Diurnal Paths of points on the Celestial Sphere

CC' Celestial Equator

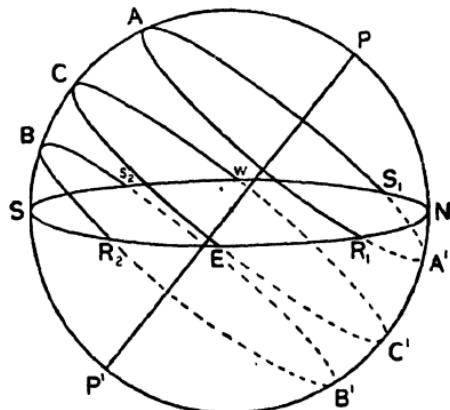
SENW Horizon

P North Celestial Pole

P' South Celestial Pole

R₁, S₁ Points of Rising and Setting of Body with northerly declination

R₂, S₂ Points of Rising and Setting of Body with southerly declination



will be evident from the diagram, are of equal length. These times, and the points on the ecliptic where the Sun then is (namely, the First Point of Aries and the First Point of Libra) are called the *equinoxes*. The First Point of Aries is the *Vernal Equinox* (about 21st March) and the First Point of Libra the *Autumnal Equinox* (about 23rd September). At these times we have equal day and night all over the Earth.

The diagram shows also that the Sun mounts much higher in the sky in the summer than in the winter. At local noon, for example, the Sun is at A near midsummer and at B near midwinter. This accounts for the greater warmth of summer as compared with winter.

Fig. 13 is drawn for a place whose latitude is near that of London. At the north pole, as we can see from Fig. 2, the Sun would appear to describe an approximately horizontal circle so long as its declination was northerly, and there would be six months of continuous sunshine. Between the autumnal and the vernal equinoxes, however, the Sun would have a southerly declination, and then it would remain below the horizon continuously for six months, so that, except for twilight, there would be six months' continuous night.* A similar state of affairs exists at the south pole, with the seasons reversed.

At the equator, as Fig. 3 shows, there is very little difference between summer and winter. The Sun describes a circle in a vertical plane every day, and at the equinoxes this circle passes through the zenith (it is then, in fact, the prime vertical). At other times it is situated at various distances on either side of the prime vertical, reaching its extreme distance at the solstices.

At places not far from either pole we get the phenomenon of the *midnight Sun*. When the Sun's declination is greater than the angular distance of the place of observation from the nearer pole, the Sun will become circumpolar, and will remain above the horizon throughout the twenty-four hours.

Let us now express the same facts in terms of the Earth's revolution round the Sun. During this revolution the Earth's axis always points in the same direction. Hence, in Fig. 14, we see that the north pole of the Earth is inclined towards the Sun when the Earth is on the left of the figure (corresponding to our northern summer) and is inclined away from the Sun at the opposite side of the orbit. In the first case, then, places in the northern hemisphere will spend most of the daily rotation in the sunlit region, and the day will be

* In a geometrical sense, that is. Actually, owing to refraction of light by our atmosphere, the Sun would appear to be above the horizon when it was actually a short distance below, so the period of continuous night would be reduced somewhat.

longer than the night, while the opposite will occur in the second case. When the Earth is midway between these positions (viz. at AE and VE in the figure) the axis will still point in the same direction, and the poles will be inclined neither towards nor away from the Sun. At these times (the equinoxes) all places on the Earth will experience twelve hours of sunlight and twelve hours of night.

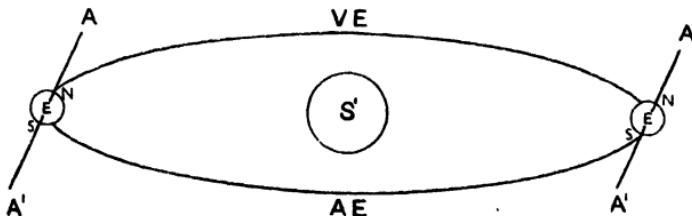


FIG. 14

Earth's Axis and Orbit

S' Sun

VE, E, AE, E Earth's Orbit

E Earth

VE Vernal Equinox

AA' Earth's Axis

AE Autumnal Equinox

The seasons, therefore, are determined by the inclination of the plane of the Earth's equator to the plane of the Earth's revolution round the Sun—or, in terms of the celestial sphere, by the inclination of the celestial equator to the ecliptic. If the equator and the ecliptic were coincident, we should have twelve hours day and twelve hours night everywhere and on every day.

EXERCISES

1. Explain the terms *sidereal time*, *mean solar time*, *equation of time*, *local time*, *standard time*. Why would it be impracticable to adopt the same clock time at the same instant for all places on the Earth's surface?
2. What, in general terms, are the differences in the appearance of the sky throughout one day at two places having

(a) the same longitude but different latitude ; (b) the same latitude but different longitude ?

3. Find the hour-angle of Aldebaran (R.A. 4 h. 32 m. 32 s.) at sidereal time 6 h. 15 m. 41 s. What is the difference between its hour-angle at Greenwich and at Athens (Long. E. $23^{\circ} 43' 15''$) at the same instant ?

4. At a certain moment a sundial at Birmingham reads 3 h. 30 m. p.m. and a watch 4 h. 40 m., the equation of time being -14 m. Find the error of the watch on British Summer Time.

5. Assuming that the Mean Sun is exactly at the First Point of Aries at mean noon on 22nd March in a certain year, calculate the sidereal time at local mean noon on 10th May of that year. If the equation of time is $+3$ m. 40 s. at that moment, find the right ascension of the Sun. The length of the year may be taken as $365\frac{1}{4}$ mean solar days.

6. If the Earth's axis were in the plane of the ecliptic, how would the alternations of day and night vary during the year at (a) one of the poles ; (b) the equator ; (c) some intermediate place ?

PART II
PROPERTIES OF MATTER

CHAPTER III

MATTER AND EXTERNAL FORCES

Molecular Theory of Matter

THE properties of matter can be understood most clearly in terms of the *molecular* or *kinetic* theory of its structure. According to this theory, every piece of matter is composed of continually moving particles called *atoms* or *molecules*, which are so small and so numerous that we cannot see them individually even with the most powerful microscope. The diameter of an atom is of the order of a hundred-millionth of a centimetre, and a molecule consists of two or more atoms in close association. For the purposes of this volume we shall not need to distinguish between atoms and molecules, and shall use the word "molecule" throughout as a general term to include both.

Between the molecules of a body, forces of attraction exist. In a solid the molecules are very close together—almost in contact, in fact—and their motions are simply oscillations about a mean position. The shapes of solids, therefore, do not change, the molecules retaining the same relative mean positions over long periods of time. In a liquid the forces of attraction are not very different in magnitude from the tendency of the molecules to separate on account of their motions. Liquids therefore change shape very easily, and conform to the shape of any vessel in which they may be placed. Nevertheless, the molecules are still very close together, and there is very little space between them. In gases, however, the spaces between the molecules are large compared with the size of the molecules themselves, and

a gas set free in a room quickly disperses and fills the room. We have familiar evidence of this in the smell of gases liberated by flowers or scent.

The student should keep in mind this conception of the structure of matter, and form the habit of thinking of any material body—of seeing it, so to speak, in the mind's eye whenever a piece of matter is thought of—as a collection of moving particles. He will then obtain a clearer understanding of the various phenomena encountered.

ELASTICITY

When an external force is applied to a body there is, in general, a change in the size and shape of the body. This

phenomenon is known as *elasticity*. The force is generally measured by its intensity per unit area of the surface over which it acts. This intensity is called the *stress*. When the force is a compression, the stress is called a *pressure*, and when it is a stretching, the stress is called a *tension*. The stress on a body may, of

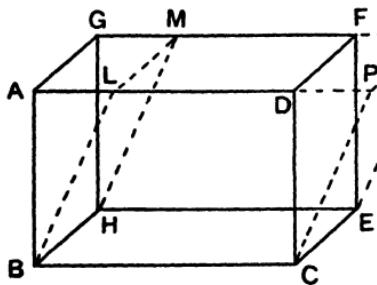


FIG. 15
Shearing Strain in Solid
ABL Angle of Shear

course, be different in different directions.

The change of shape or size of a body resulting from a stress is called a *strain*. If the shape remains the same, and only the size changes, the strain is called a *uniform dilatation* if the body is enlarged, or a *uniform compression* if it is diminished. On the other hand, if the size remains the same and only the shape changes, the strain is called a *shear*. In general, therefore, a stress can be regarded as made up of a uniform dilatation or compression and a shear.

A shear is often most conveniently regarded as a displacement of successive layers of the substance by an amount proportional to the distance of the layer from a fixed, undisplaced surface. Thus, in Fig. 15, where the continuous lines represent the unsheared and the dotted lines the sheared positions of a block of material, the horizontal displacement of a layer is proportional to its distance from the face BCEH, and the stress is a "tangential" force parallel to this face. The angle ABL in this case is called the *angle of shear*. It is clear from elementary geometry that the block is not changed in volume by the displacement.

Hooke's Law

The relation between stress and strain must be determined by observation. Provided the stress does not exceed a certain value, which depends on the material, this relation is expressed by *Hooke's Law*, viz. *Ut tensio sic vis*, or *The strain is proportional to the stress*. This means, for example, that if a small force produces a certain change of size or shape, twice that force will produce twice the change.

Moduli of Elasticity

The ratio of stress to strain, which by Hooke's Law is constant for moderate stresses, is called a *modulus of elasticity*. For a uniform compression or dilatation it is called the *bulk modulus*, and for a shear it is called the *modulus of rigidity*. A special modulus is defined for the case of a wire stretched longitudinally, because of its frequent occurrence in important problems. Here there is a change of both size and shape of the wire, but we are interested only in the change of length, and the corresponding modulus is called *Young's modulus*. These three moduli are defined more precisely as follows. The bulk modulus, k , is the applied force (compression or tension) per unit area divided by the change of volume per unit volume. The modulus of rigidity, n , is the tangential force per unit area, divided by the angle of shear. Young's modulus, q , is the force

per unit cross-section of the wire divided by the increase of length per unit length.

Clearly, the greater the value of a modulus, the more resistant is the material to deformation. The moduli of elasticity will thus vary enormously for matter in different states. Liquids, for example, have extremely small moduli of rigidity, whereas solids, as a rule, have high rigidity moduli. We shall in this section treat both solids and liquids, but owing to the close relation between the mechanical and thermal properties of gases, it will be more convenient to treat the elasticity of the gaseous state in the section on heat.

Since the bulk modulus and the modulus of rigidity measure all the elastic changes to which a body can be subjected, Young's modulus must be expressible in terms of them. It can be shown that the relation is :

$$\left. \begin{aligned} q &= \frac{9nk}{3k + n} \\ \text{or} \quad \frac{1}{q} &= \frac{1}{3n} + \frac{1}{9k} \end{aligned} \right\} \quad \dots \quad \dots \quad \dots \quad \dots \quad (3.1)$$

Of course, when a wire is stretched there is a contraction in its lateral dimensions, of which no account is taken in Young's modulus, because we are often not interested in it. Sometimes, however, it is important. The ratio of the lateral to the longitudinal strain of a bar or wire is called *Poisson's ratio*. Thus, if a wire, originally of length l and radius r , is stretched so that its length increases by dl and its radius decreases by dr , Poisson's ratio, σ , is given by

$$\sigma = \frac{dr/r}{dl/l} \quad \dots \quad \dots \quad \dots \quad \dots \quad (3.2)$$

In terms of the bulk and rigidity moduli,

$$\sigma = \frac{3k - 2n}{2(3k + n)} \quad \dots \quad \dots \quad \dots \quad \dots \quad (3.3)$$

a fraction which must be less than $\frac{1}{2}$ and is usually not far from $\frac{1}{4}$.

The moduli of elasticity vary considerably from one specimen to another, even of the same material, depending on its physical condition. For solids they are of the order of 10^{11} . There are various methods for determining them. We shall describe one method for Young's modulus and one for the modulus of rigidity. Before doing so, however, it is necessary to see the limits within which Hooke's law is true, for outside those limits the moduli lose their significance.

The Elastic Limit

Let a wire hang vertically and be stretched by weights which can be gradually added to a scale-pan suspended from its lower end. If the length of the wire be measured at each change of weight, it will be found that at first Hooke's law is obeyed and, further, if at any stage the weights are removed, the wire returns after a time to its former length. As the load increases, however, a point is reached when the extension suddenly increases at an abnormally high rate, and when the weights are removed the rod does not recover its former state. This point is called the *elastic limit*. If the load be continued after the elastic limit is passed, the changes which occur are somewhat complex, until ultimately the *breaking-point* is reached, when the wire breaks. Somewhat similar changes occur with shearing stresses.

We shall not consider stresses which take the material beyond the elastic limit, but it must be noted that, with even small forces, there is great variety with different materials in the time taken to recover the original condition when the stress is removed. Glass, for example, recovers slowly, and glass vessels, such as thermometers, which are subjected to changes of temperature, do not return readily to their true zero points. Quartz, on the other hand, recovers its original state almost immediately, and on this account quartz fibres are specially suitable for suspending galvanometer coils and other systems in measuring instruments. Again, even the

same specimen of material will usually not show the same behaviour when two successive loading experiments are made with it. It takes several loadings and unloadings to bring the substance to a condition in which its behaviour can be depended on.

Measurement of Young's Modulus

A convenient and straightforward method for determining Young's modulus is to hang a scale-pan from a vertically suspended wire carrying, near its bottom, a vernier which can be read against a fixed scale. As weights are gradually added to the scale-pan, the vernier sinks lower and lower, and the scale indicates its successive positions. The

stress is then the weight in the pan (together, of course, with the weight of the pan itself) divided by the area of cross-section of the wire; and the strain is the lowering of the vernier divided by the original distance of the vernier from the top of the wire. Another method will be mentioned later (p. 56) involving the bending of beams.

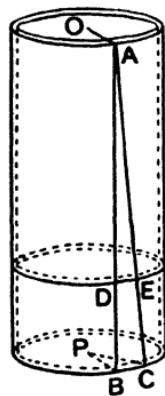


FIG. 16

Shearing of thin tube
 $OA, PB, PC = r$
 $\angle BPC = \theta$
 It is clear from the figure that the displacement of any point along AC from its original position is proportional to

Measurement of Rigidity

To understand the method of determining the modulus of rigidity, let us first consider a thin tube, of circular section, of which one end is fixed and the other twisted through an angle θ by a couple applied in a plane perpendicular to the axis of the tube. Thus, in Fig. 16, the lower end is twisted through the angle $\theta = BPC$ and the line AB has become AC. Let the radius of the tube (internal or external, since the tube is assumed very thin) be r , and the thickness of its material, dr . It is clear from the figure that the displacement of any point along AC from its original position is proportional to

its distance from A : thus, the displacement of C from B is to the displacement of E from D as AB is to AD. Hence

$$BC - DE = BC \left(1 - \frac{AD}{AB}\right) = BC \frac{DB}{AB} \dots \dots \quad (3.4)$$

Let us call this α .

The portion of material between the horizontal planes through DE and BC is thus in a state of shear, the layer at the bottom being displaced by α with respect to a layer at a distance DB above it. The angle of shear is therefore $\frac{\alpha}{DB}$, i.e. $\frac{BC}{AB}$. Now $BC = r\theta$, since $\theta = BPC$: hence the angle of shear is $\frac{r\theta}{AB}$. Since the displacement of a point is proportional to its distance from the fixed end, $\frac{\theta}{AB}$ is constant, wherever B is chosen : it is the angle of twist per unit length of the tube. Let this angle be ϕ . Then the angle of shear is $r\phi$; ϕ , of course, being measured in radians.

Now, from the definition of the rigidity modulus,

$$n = \frac{T}{r\phi} \dots \dots \dots \dots \quad (3.5)$$

where T is the tangential stress in the tube. Hence, to find n we must now find T . If the tube is in equilibrium, the moment about its axis of the force corresponding to T must be equal to the moment, M , of the external couple which causes the twist. The force corresponding to T is T multiplied by the area of cross-section of the material of the tube ; i.e. $T \times 2\pi r dr$. The moment of this force is therefore $2\pi r^2 dr T$. Hence :

$$\begin{aligned} M &= 2\pi r^2 dr T \\ &= 2\pi r^3 \phi n dr, \text{ from (3.5)} \dots \dots \quad (3.6) \end{aligned}$$

To apply this result experimentally we must arrange to keep one end of the tube fixed, and twist the other by a couple ^{*}

which can be accurately measured. This may be done by means of threads passing tangentially from the tube over pulleys, and carrying scale-pans in which known weights can be placed. The angle of twist can be measured by a light pointer attached to the tube and moving over a horizontal scale of degrees. If this pointer is at a distance l from the fixed end, and its reading is β , then

$$\phi = \beta/l. \quad . \quad . \quad . \quad . \quad . \quad (3.7)$$

The radius and thickness of the tube can be measured by ordinary methods—the former being determined as accurately as possible, since it occurs to the third power, and any error is therefore magnified in the calculation.

It is easy to apply the same principle to a cylindrical rod instead of a tube. For such a rod may be regarded as composed of a series of concentric tubes of gradually increasing radii, to each of which the above calculation can be applied. The couple twisting the rod must then be equal to the sum of the couples necessary to twist this series of tubes, and this is obtained by integrating equation (3.6) from $r = 0$ to $r = R$, the external radius. We then obtain

$$M = 2\pi\phi n \int_0^R r^3 dr = \frac{\pi}{2} \phi n R^4. \quad . \quad . \quad . \quad . \quad (3.8)$$

The experimental arrangements are the same as with the tube, but we avoid the error arising from irregularities in dr .

Bending of Beams

The bending of beams affords a good example of the application of the principles of elasticity. Suppose a beam of some solid material is supported on knife-edges near its two ends (Fig. 17) and loaded by a weight W at the middle. The beam will be bent as shown, and if the weight is not too great it will remain in equilibrium in that state. The lower edge of the beam is then stretched, and the upper edge compressed, so that between the two there is a layer which is neither ex-

tended nor contracted : this is called the *neutral surface* of the beam. It can be shown that this surface is very nearly an arc of a circle.

It is clear that the amount by which a given bar is bent by a given weight when in equilibrium will depend on Young's modulus, q . For the couples formed by the weight and the two supports must be balanced by a couple consisting of a stretching and compressing force, P (Fig. 17), in order that equilibrium shall exist, and the bending will depend on the strain which this couple sets up in the bar.

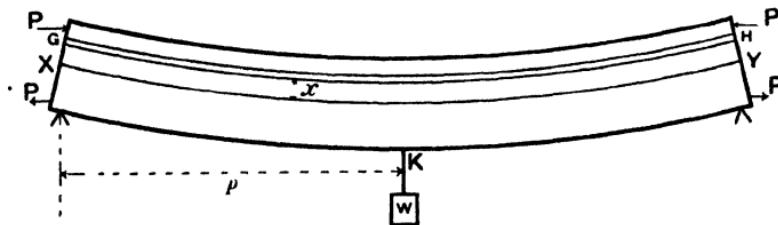


FIG. 17

Bending of Beam, supported at ends, by central weight

XY Neutral Surface

GH Layer distant x from XY

Consider a small element of the bar in the form of a thin strip extending along its length—GH in Fig. 17—distant x from the neutral surface, XY, and having an end area a . Then, if F be the value of P per unit area, the end force on this element will be Fa , and the moment of this force about XY will be Fax . Suppose this element is compressed (or stretched, if it is below XY) by an amount y . Then, if L is its original length (*i.e.* the permanent length of the neutral surface), the strain is $\frac{y}{L}$, so that the stress, F , is $q\frac{y}{L}$. Hence the moment about XY of the end force on the element is $q\frac{y}{L}ax$.

The total moment of the couple P is obtained by adding together all such quantities as this for all the elements which

make up the beam. Let this sum be $\Sigma q \frac{y}{L} ax$,* summed over the area of one end of the beam. This must be balanced by the moment of the applied couple, which, for the corresponding half of the bar, is $\frac{W}{2} \times p$ (see Fig. 17). Hence

$$\frac{Wp}{2} = \Sigma q \frac{y}{L} ax. \quad \dots \dots \dots \quad (3.9)$$

We can simplify this by expressing the strain, $\frac{y}{L}$, in terms of the distance, x , from the neutral surface. For, if r be the radius of the circle into which, approximately, the neutral surface is bent, we see that

$$\frac{\text{arc GH}}{\text{arc XY}} = \frac{r - x}{r} \quad \dots \dots \dots \quad (3.10)$$

whence

$$\frac{\text{XY} - \text{GH}}{\text{XY}} \left(= \frac{y}{L} \right) = \frac{x}{r} \quad \dots \dots \dots \quad (3.11)$$

Hence

$$\frac{Wp}{2} = \Sigma q a \frac{x^2}{r} = \frac{q}{r} \Sigma a x^2 \quad \dots \dots \dots \quad (3.12)$$

since q and r are the same for all elements.

The quantity $\Sigma a x^2$ is known as the "moment of inertia of the cross-section" of the rod ; its value depends on the shape of the cross-section, and can be calculated when this shape is known. Let us denote it by I . Then

$$\frac{Wp}{2} = \frac{qI}{r} \quad \dots \dots \dots \quad (3.13)$$

whence Young's modulus, q , is given by

$$q = \frac{Wpr}{2I} \quad \dots \dots \dots \quad (3.14)$$

* The symbol Σ (Greek " sigma ") is used to indicate the sum of all such quantities as that placed after it. Thus $\Sigma q \frac{y}{L} ax$ means the sum of all the quantities $q \frac{y}{L} ax$ for all the elements of area, a , which make up the area of the end of the beam.

We can determine r approximately by measuring the depression of the centre, K , of the lower surface of the beam below the points of support, on the assumption that the beam is bent into a circular arc. If this depression be h , and the depth of the beam be d , the radius of the assumed circle is $r + \frac{d}{2}$, and we have

$$h \left[2 \left(r + \frac{d}{2} \right) - h \right] = p^2 \quad \dots \quad (3 \cdot 15)$$

whence, neglecting the small quantity, $dh - h^2$,

$$r = \frac{p^2}{2h} \quad \dots \quad (3 \cdot 16)$$

The modulus, q , is therefore given by

$$q = \frac{Wp^3}{4hI} \quad \dots \quad (3 \cdot 17)$$

or, since p is approximately $\frac{L}{2}$, when the beam is supported at its ends,

$$q = \frac{WL^3}{32hI} \quad \dots \quad (3 \cdot 18)$$

The exact value, taking into account the true curvature of the beam which is not exactly that of a circle, is

$$q = \frac{WL^3}{48hI} \quad \dots \quad (3 \cdot 19)$$

an equation which affords a very convenient method for determining q .

It might be mentioned that, when the cross-section of the beam is rectangular, of breadth b and depth d , $I = \frac{bd^3}{12}$, so that, in this case,

$$q = \frac{WL^3}{4hbd^3} \quad \dots \quad (3 \cdot 20)$$

Elasticity of Liquids

It is obvious that liquids offer very little resistance to change of shape, but they offer considerable resistance to change of size. Liquids are, it is true, more easily compressible than solids (water, for example, has at ordinary temperatures a bulk modulus of about $2 \cdot 10^{10}$ —i.e. the change of volume per unit volume under a pressure of one atmosphere is equal to about $5 \cdot 10^{-5}$ c.c.), but the difficulty in compressing them is great enough for the belief to have been held until the middle of the eighteenth century that they were incompressible. The compressibility of liquids, nevertheless, is not without important consequences. It has been calculated that if sea water were incompressible, the sea-level would be several feet higher than it is, with the result that hundreds of thousands of square miles of land surface would be submerged.

Measurements of the compressibility of liquids are complicated by the fact that the vessels in which they must be contained also change volume under compression, and this change must be determined separately and allowed for. We have a similar complication in measuring the expansion of liquids with heat (p. 109), a subject which obviously has close relations with elasticity.

It might be thought that liquids would be unable to support a tension of any strength, but this is not so. It is experimentally difficult to establish conditions in which the tensile strength of a liquid can be determined without the liquid changing shape and running away, but when this difficulty is overcome it is found that liquids can withstand very great pulls. It is necessary to remove all dissolved air from the liquid under test before the experiments are made, otherwise the liquid will break very easily. Definite values for the moduli of elasticity have not been obtained, but it has been established that liquids can endure a pull of many atmospheres without fracture.

Elasticity on the Molecular Theory

The more general elastic properties of solids and liquids are easily understood in terms of the molecular constitution of matter. Evidently we must suppose that there are forces between the molecules which resist tensions and equally resist compressions. Hence the molecules must stand in positions of extremely stable equilibrium with respect to one another. This does not apply to gases, however, as we shall see, but in solids and liquids we have reason to think that the molecules are practically in contact, so that, however great a force of compression we apply, a substance cannot occupy a much smaller space than it normally does.

The existence of an elastic limit suggests that when the molecules have been pulled apart a certain distance, the forces of attraction are reduced, and it is then easier to produce additional enlargement. This is somewhat analogous to the phenomenon of melting, and indeed a solid beyond the elastic limit shows a tendency toward acquiring the properties of the liquid state. This, however, must be regarded only as a very general explanation of the facts. The details are much complicated, particularly with metals, by the crystalline form which the substances assume.

EXERCISES

1. Give an account of the molecular theory of the structure of matter, explaining how it accounts for the main differences in the properties of solids, liquids, and gases.
2. State Hooke's law. A model barrage balloon is supported by 6 cables, each of cross-sectional area 0.1 sq. cm., at a height of 100 feet. The cables are found to be stretched by a length of 4 inches. Taking Young's modulus as 2×10^{12} dynes per sq. cm., find the additional weight which the balloon could support.

3. A straight rod, 1 metre in length, and with a circular cross-section 1 sq. cm. in area, is hung vertically with the upper end fixed. At the lower end two horizontal strings issue tangentially from opposite ends of a diameter of the circular end of the rod, and each, after passing over a pulley, supports a weight of 20 kilograms. The lower end of the rod is then found to be twisted through an angle of 1° . Find the modulus of rigidity of the rod.

CHAPTER IV

MATTER AND INTERNAL FORCES

COHESION

THE existence of internal forces in matter—at least in solids and liquids—has already been shown by the fact that bodies resist the action of external forces on them. We have just seen that these internal forces are very great. The property of remaining whole with which they endow bodies is known as *cohesion*. Experimental investigations of cohesion can best be carried out by submitting matter to the influence of external forces, in the manner considered in the last chapter, or by the application of heat, and the cohesive forces are indicated by the moduli of elasticity or the coefficients of expansion (see p. 106).

SURFACE TENSION

There is one important set of phenomena, however, in which cohesion manifests itself directly—namely, phenomena

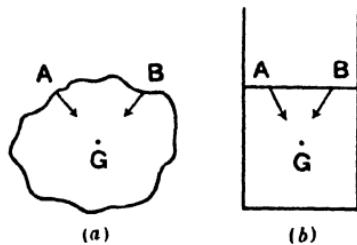


FIG. 18
Molecular attraction in
(a) solid (b) liquid
G Centre of mass
A, B Points on surface

connected with the surfaces of bodies. To see how this comes about, consider the solid and liquid represented in Fig. 18, in which G represents the centre of mass, and A and B are two points on the surface. Now a molecule in the neighbour-

hood of G will be surrounded by other molecules, all of which attract it, so that it is pulled equally on all sides and therefore has no tendency to move in any particular direction. If, as we have reason to believe, the force of attraction diminishes rapidly as the distance between the molecules increases, so that a molecule is effectively attracted only by its immediate neighbours, then the same is true for all the molecules in the solid or liquid except those on the surface. A surface molecule, however, is attracted only by those inside the body, for there are none outside (the attraction of the molecules of the air or other matter surrounding the body is in general of a different order of magnitude, and will be left out of account for the present), and the resultant attraction on it is in the direction of the centre of mass G. The molecules at A and B, therefore, are pulled in the directions shown by the arrows, *i.e.* they have a tendency to leave the surface for the interior, and also a tendency to approach one another on the surface, as we may see by resolving the forces along and at right angles to the surface. These tendencies indicate that, if no other influences are at work, a body will tend to have as small a surface as possible, and its surface will act as though it were in a state of tension. This phenomenon is known, in fact, as *surface tension*.

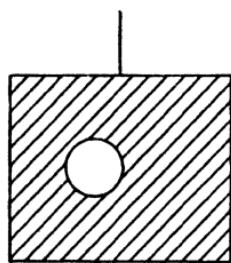


FIG. 19

Tension in surface of
Soap Film

surface molecule, however, is attracted only by those inside the body, for there are none outside (the attraction of the molecules of the air or other matter surrounding the body is in general of a different order of magnitude, and will be left out of account for the present), and the resultant attraction on it is in the direction of the centre of mass G. The molecules at A and B, therefore, are pulled in the directions shown by the arrows, *i.e.* they have a tendency to leave the surface for the interior, and also a tendency to approach one another on the surface, as we may see by resolving the forces along and at right angles to the surface. These tendencies indicate that, if no other influences are at work, a body will tend to have as small a surface as possible, and its surface will act as though it were in a state of tension. This phenomenon is known, in fact, as *surface tension*.

It is not easy to get direct evidence that the surface of a solid is in a state of tension because the great rigidity of solids keeps their shape constant. We shall therefore restrict our attention to liquids. These show the phenomenon quite well in a variety of ways, of which the following is one simple example. Dip a loop of wire into a soap solution so as to fill it with a thin film, and drop on this film a small loop of thread. If the film inside the thread is then pierced, the thread is immediately drawn into a circle (Fig. 19). The

remaining film tries to contract, and so pulls the thread into the form which gives the film the smallest surface.

Surface Energy

Another way of looking at the phenomenon of surface tension is of interest. Any mechanical system tends to come into the position in which its potential energy is a minimum. Thus, bodies free to move fall to the ground ; and the surface of a liquid is horizontal because it cannot be tilted without raising the centre of gravity of the liquid and so increasing the potential energy. Now we have seen that when a molecule moves from the interior to the surface of a liquid, work must be done on it because it is moving against the direction of the resultant force. A surface molecule, therefore, has greater potential energy than an interior molecule. Consequently the substance tends to get into the state in which it has the smallest number of surface molecules—*i.e.* that in which it has the smallest surface.

It frequently happens that this tendency is overcome by other influences. Thus, if water is spilt on the ground, it does not collect into a sphere but spreads out. In this case the tendency towards minimum gravitational potential (lowering of the centre of gravity) overcomes that towards minimum surface energy, and the liquid takes the form in which the *total* potential energy is least. If the amount of water is very small, however, so that very little loss of gravitational energy is achieved by spreading, surface tension becomes dominant, and the liquid collects into a sphere. It may be calculated that water-drops less than two or three millimetres in radius will remain spherical and will not spread.

Definition of Surface Tension

A formal definition of surface tension is obtained as follows. Suppose a film of liquid (*e.g.* soap solution) fills a framework ABCD (Fig. 20) of which the side CD can be moved to and fro. The surface tension of the film will tend to pull CD

towards AB, and a force P must therefore be applied to keep CD at rest. The surface tension, T , of the film is then defined by

$$T = \frac{P}{2CD} \dots \dots \dots \quad (4.1)$$

This means that the surface tension is the pull per centimetre exerted by the film, the factor 2 in the denominator appearing because the film has two sides, *i.e.* two surfaces. The value of T for pure water at 15°C . is about 74 dynes per cm.

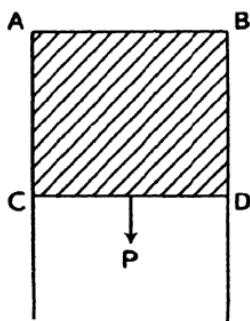


FIG. 20
Soap Film in
Equilibrium

We can now easily evaluate the increase in surface energy caused by an enlargement of the surface. For, suppose CD is pulled away from AB through a distance x . The work done is then Px , *i.e.* $2Tx \cdot CD$, and this must be equal to the energy acquired by the film. If we neglect any change of temperature that might accompany the stretching, the energy acquired by the film is entirely surface energy, and since the area is increased by $2x \cdot CD$, we see that T is equal to the surface energy acquired per unit area when the surface is enlarged. This is a very useful way of thinking of surface tension.

A commonly observed effect of surface tension is that a liquid behaves as though it were enclosed within a stretched skin. Thus, a drop of water hanging from the end of a glass rod takes the same form as though it were enclosed within a flexible bag. It is, in fact, supported, not by the attraction of the rod but by the tension in its surface. If the drop fell, its surface would be enlarged and its potential energy therefore increased, so it remains suspended.

The surface of a liquid (see Fig. 18(b), for example) is not only that boundary adjacent to the air but also the area of contact of the liquid with its containing vessel. There will

be in general a tension in this area also, but this will not be the same as the tension in the air boundary, for a molecule here will be attracted appreciably by the neighbouring molecules of the solid vessel, whereas the air molecules, moving freely, will mostly be too distant to have any appreciable effect. It follows that a liquid in contact partly with a solid and partly with air will adjust the relative areas of contact so that the total surface energy (in so far as gravitational considerations are not predominant, as they usually are) is a minimum.

Angle of Contact

We can show that this condition will result in the establishment of a definite *angle of contact* between the liquid and the solid (usually the containing vessel). Suppose EFGH (Fig. 21) represents one face of a rectangular vessel partly filled with liquid whose air surface is represented by LMNP. We have then three surfaces to consider, namely :

EFLM—solid/air surface

LMNP—liquid/air surface

and LMGH—solid/liquid surface.

Let the respective surface tensions be T_{sa} , T_{li} , T_{sl} . The condition of equilibrium (apart, of course, from the predominant gravitational necessity that the liquid shall, in the main, settle in the lowest part of the vessel), so far as the common line of contact, LM, is concerned, is that the total surface energy shall be a minimum; and this means that if the line be displaced upwards, say, by an infinitesimal amount, the surface energy shall not change.*

Let us, then, consider a small change in the surface energy

* The definition of a configuration giving a minimum or maximum value of a quantity is that configuration in which an infinitesimal displacement causes no change in the value. Thus a pendulum hanging at rest has a minimum potential energy, and an infinitesimal displacement of the bob is in a horizontal direction, giving no change of potential energy.

resulting from a small change in the angle of contact between the liquid and the solid. This is shown in Fig. 21, where the new position of LM is shown as L'M'. By this change the solid/liquid surface is enlarged by the area $L'M'ML$ ($= A$, say), and the solid/air surface is diminished by the same amount. The net increase of surface energy on account of these two changes is therefore $A (T_{sl} - T_{sa})$. The liquid/air surface, however, is increased by $L'M'VW - LMVW$, and by drawing perpendiculars from L and M on the surface

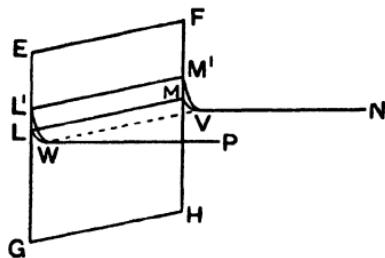


FIG. 21

Angle of Contact of Liquid and Solid.

$L'M'VW$, we see that this difference is $L'M'ML \cos GL'W$, *i.e.* $A \cos \theta$, where $\theta = GL'W = GLW$, since we are considering only an infinitesimal change. The increase of energy corresponding to the enlargement of the liquid/air surface is therefore $T_{la} \cdot A \cos \theta$, so that the total increase of energy is $A (T_{sl} - T_{sa} + T_{la} \cos \theta)$.

Now, as we have seen, if the original position is one of equilibrium, this increase must be zero. Hence

$$\cos \theta = \frac{T_{sa} - T_{sl}}{T_{la}}. \quad \dots \quad (4.2)$$

The liquid will therefore adjust itself so that an angle θ , satisfying this relation, is established. If the only surface tension concerned in the matter had been that of the liquid/air surface, θ would have been 90° , for that would have given the smallest surface. The fact that θ is not 90° in general may therefore be taken as evidence that there actually is a tension at the other surfaces.

It is clear that θ will be less or greater than a right angle according as T_{sa} is greater or less than T_{sl} . Most liquids

and solids are such that the former relation holds, and Fig. 21 represents the equilibrium position in this case. It is assumed, however, in deriving this result that the liquid does not wet the tube. If it does, as in the ordinary case of water in a glass vessel, then the liquid still rises at its edge, but the angle of contact diminishes to zero, the side of the vessel becoming tangential to the curved liquid surface. With mercury, however, T_s is greater than T_{sa} , so that θ is greater than a right angle. We thus get the very familiar appearance of the surface of mercury in a barometer tube, with the convex side upwards.

A very simple method of finding the angle of contact of, say, mercury and glass may be mentioned. Hold a plane sheet of glass vertical in a vessel of mercury, and gradually

tilt it about a horizontal axis. At a certain angle (ACD in Fig. 22) the mercury will meet the glass without any curvature of the surface. This angle is the supplement of the angle of contact, θ .

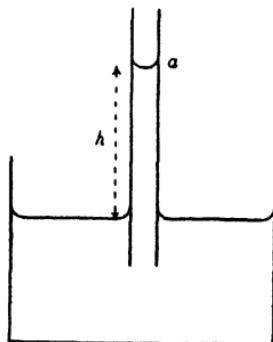


FIG. 23

Rise of Liquid in Capillary Tube

the height of the liquid in the capillary tube above the level of liquid outside when equilibrium is established. Then, as before, the potential energy will be unchanged if the liquid is raised an infinitesimal distance in the tube. Let it be

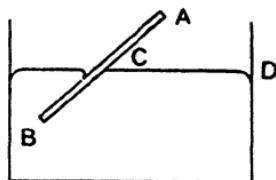


FIG. 22

Determination of Angle of Contact of Mercury and Glass

Liquids in Capillary Tubes

The well-known fact that a liquid rises or sinks in a capillary tube can now be understood. Fig. 23 represents the case of a liquid such as water, for which θ is less than a right angle. Let the radius of the tube be r , and let h be

raised such a distance, which we will denote by dh . Then the liquid/air surface remains unchanged, but the solid/liquid surface is increased, and the solid/air surface decreased, by an area $2\pi r dh$, so that the total increase of surface energy is $2\pi r dh(T_{sa} - T_{sl})$.

But there is also an increase in gravitational potential energy, since liquid has been raised in the tube. To evaluate this increase we may consider the meniscus above the horizontal tangent at a (of volume v , say) to be raised through a height dh , and the space thus left to be filled by raising a volume of liquid, $\pi r^2 dh$, through a height h . If the density of the liquid be ρ , then the work done in raising the meniscus is $v \rho g dh$, and that done in raising the small cylinder of liquid is $\pi r^2 dh \rho g h$. The total work done, which is the increase of gravitational potential energy, is therefore $\rho g dh(v + \pi r^2 h)$.

The complete increase of potential energy, surface and gravitational, is thus

$$2\pi r dh(T_{sa} - T_{sl}) + \rho g dh(v + \pi r^2 h)$$

and this, as we have seen, must vanish. Hence we must have

$$T_{sa} - T_{sl} = \frac{\rho g(v + \pi r^2 h)}{2\pi r} \quad \dots \quad (4.3)$$

which, from (4.2), is equivalent to

$$\left. \begin{aligned} T_{sa} \cos \theta &= \rho g \left(\frac{v}{2\pi r} + \frac{rh}{2} \right) \\ \text{or } h &= \frac{2T_{sa} \cos \theta}{\rho gr} - \frac{v}{\pi r^2} \end{aligned} \right\} \quad \dots \quad (4.4)$$

If, then, we take a series of gradually narrowing tubes, h will gradually increase, since the second term increases more rapidly than the first as r diminishes. This accords with experience. (On this account the phenomenon of surface tension is sometimes called *capillarity*.) In the common case of a

liquid which wets the tube, $\theta = 0$ and $\cos \theta = 1$. We can then evaluate v for narrow tubes in which the meniscus may be regarded as a hemisphere, for it is the difference between the volume of a hemisphere and the cylinder which circumscribes it, viz. $\frac{\pi r^3}{3}$. Equation (4.4) then becomes

$$h = \frac{2T_{\text{L}}}{\rho gr} - \frac{r}{3} \quad \dots \quad \dots \quad \dots \quad \dots \quad (4.5)$$

From (4.4) we see also that when $\cos \theta$ is negative, h is negative, *i.e.* the liquid sinks in the tube. This occurs with mercury.

It should be mentioned that the angle of contact can be made to vary considerably if the liquid and solid are not very clean. It is necessary, therefore, to clean them as thoroughly as possible if this method is used for finding the value of T_{L} .

Pressure in Bubbles

Yet another consequence of the existence of surface tension is the fact that a bubble of air or vapour in a liquid will be at a higher pressure than that existing in the liquid just outside. For if the pressures are initially the same, the tendency for the surface to contract will be unresisted, and the bubble will shrink until its tendency to do so is balanced by the increased internal pressure resulting from the reduced volume of the gas (see p. 121).

In the case of a spherical bubble we can easily calculate what the excess of pressure must be. For, in Fig. 24, if we imagine the bubble to be divided into two hemispheres by a horizontal plane, the internal force across this plane, tending to break the bubble in two, is $\pi r^2 p$, where r is the radius of the bubble and p is the excess of pressure over the pressure outside which

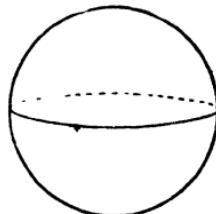


FIG. 24

Spherical Bubble

tends to keep the sphere intact. This force is balanced by that due to surface tension, which is $2\pi rT$. Hence

$$p = \frac{2T}{r} \dots \dots \dots \dots \quad (4.6)$$

It is assumed here that the bubble is entirely surrounded by liquid. If we have a thin soap bubble, say, with air outside, then, since there are two surfaces, the surface tension can sustain twice this excess of pressure and we have

$$p = \frac{4T}{r} \dots \dots \dots \dots \quad (4.7)$$

It will be noticed that as r diminishes, the excess of pressure increases. A very small bubble, therefore, would tend strongly to collapse, in order to reach the pressure necessary to resist the surface contraction.

The subject of surface tension is a very large and important one. The phenomena discussed here are to be regarded as examples of some of the more important of its aspects. It can be understood, however, that since every liquid must have a surface across which there will be a tension, there is no problem concerning the behaviour of liquids in which it does not play some part.

EXERCISES

1. Explain why, on the molecular theory of matter, you would expect a stress to appear on the surface of a liquid.
2. The surface of a liquid contained in a solid vessel is curved at the edges. Explain why this is so, and show how the direction of curvature is related to the rising or sinking of the liquid in a capillary tube.
3. The liquid film shown in Fig. 20 is in equilibrium when

a weight of 0.75 gm. is hung from CD. If CD is 5 cm. long, what is the surface tension of the liquid ?

4. Taking the surface tension of water to be 74 dynes per cm., calculate the diameter of a capillary tube in which water rises to a height of 1 cm. above the outside level.

5. The pressure in a liquid at a depth h below the surface is equal to the pressure of the atmosphere plus gph , where ρ is the density of the liquid. If the pressure is the same inside two spherical gas bubbles in equilibrium in the liquid, one of which is 1 cm. lower than the other, and the radius of the upper one is 0.1 mm., find the radius of the other.

CHAPTER V

INTERACTION BETWEEN MATTER AND MATTER

Types of Interaction

Two pieces of matter may interact with one another either across empty space or through actual contact. The former type of interaction is exemplified by gravitational attraction (which, apart from the fact that bodies have weight, is more important in astronomy than in physics), by electric and magnetic forces (considered in the second volume of this book), and by exchange of heat—or, more generally, radiant energy (considered in Chap. XII.). We restrict our attention here to the effects of contact of matter with matter.

When two pieces of matter meet one another, various things may happen. We leave out of account, of course, the occurrence of chemical combination, which is a large and distinct subject, and consider only physical changes in the narrow sense. On meeting, then, two bodies may strike and rebound ; this we consider under the title of *impact*. They may interpenetrate ; this gives rise to the phenomena of *solution* and *diffusion*, and, in the special case of the passage of one body completely through the other, to the phenomenon of *osmosis*. Finally, they may resist one another's motion without either rebounding or interpenetrating. We have then the phenomena of *friction* and *viscosity*. Friction between solids is usually considered as a department of mechanics, and we shall therefore leave it out of consideration, but we shall deal with the essentially similar phenomenon of viscosity which characterizes the behaviour of fluids.

IMPACT

The phenomena of impact occur chiefly with solids. It is possible to make a solid rebound from a liquid, as in the children's pastime of throwing pebbles horizontally on to the surface of the sea, and so causing "ducks and drakes." Liquids also rebound from the surfaces of liquids—a process familiar to us as splashing. We shall, however, ignore these comparatively unimportant phenomena and consider only the impact of solids with solids.

This may be investigated by suspending two balls from a fixed support and allowing them to collide with one another as they swing in the same plane (Fig. 25). If either ball be displaced a certain distance and then released, its velocity on returning to its lowest position will depend only on the amount of the displacement, and can be calculated therefrom—the resistance of the air being allowed for. We can thus make the balls execute a head-on collision, when each has any velocity we wish within the limits set by the dimensions of the apparatus. In general the balls will rebound after collision, and the initial velocities of rebound can be determined from the distance through which the balls travel thereafter.

By such experiments it is found that if u_1, v_1 be the velocities immediately before, and u_2, v_2 the velocities immediately after collision (u and v have opposite signs, both before and after, if the velocities are in opposite directions), the following relation holds for a given pair of balls :

$$u_2 - v_2 = e(v_1 - u_1) \quad \dots \quad (5.1)$$

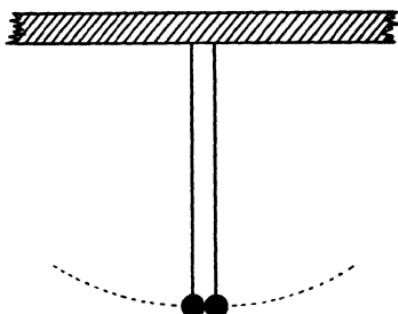


FIG. 25

Apparatus for studying impact

where e is a constant depending on the material of which the balls are made. e is called the *coefficient of restitution*. It is independent of the masses of the balls, and, for moderate velocities, it is independent of the initial velocities also, though it tends to decrease as the velocities become very large.

It will be noticed that equation (5.1) does not tell us what the final velocities will be ; it gives us only their difference. To determine the individual velocities we must know the masses of the balls. If these are M and m , respectively, and we assume that the law of conservation of momentum is observed in the impact, then we have also the equation :

$$Mu_1 + mv_1 = Mu_2 + mv_2 \quad . . . \quad (5.2)$$

and from the two equations, (5.1) and (5.2), we can calculate the final velocities. We easily obtain :

$$\left. \begin{aligned} u_2 &= \frac{mv_1 + Mu_1}{m + M} + e \frac{m(v_1 - u_1)}{m + M} \\ v_2 &= \frac{mv_1 + Mu_1}{m + M} - e \frac{M(v_1 - u_1)}{m + M} \end{aligned} \right\} \quad . . . \quad (5.3)$$

It follows that, in general, there is a loss of kinetic energy on impact. For the initial kinetic energy is

$$\frac{1}{2}Mu_1^2 + \frac{1}{2}mv_1^2$$

and equations (5.3) show that the final kinetic energy (which is, of course, $\frac{1}{2}Mu_2^2 + \frac{1}{2}mv_2^2$) is equal to

$$\frac{1}{2}Mu_1^2 + \frac{1}{2}mv_1^2 - \frac{1}{2}(1 - e^2) \frac{Mm}{M + m} (v_1 - u_1)^2$$

Now e is never greater than 1, so that, unless $e = 1$, the final is less than the initial kinetic energy. The lost energy is transformed into heat, or energy of vibration of the system, or sound, and passes beyond our control.

Perfect Elasticity

This transformation of energy is not direct. The first thing that happens on impact is that the balls are flattened at the point of contact, and come instantaneously to a common velocity. At this instant some of the initial kinetic energy exists as strain energy in the balls. It depends on the material of which the balls are made in what measure this strain energy is converted back into kinetic energy again. If e is equal to 1 (a condition which is never completely realized in practice) the balls are said to be *perfectly elastic*. If it is equal to 0 (in which case the balls remain together after impact) they are said to be *perfectly inelastic*. Usually e is a fraction, the balls partly converting the strain energy into kinetic energy by repelling one another, and partly transmitting it through the system as vibration or heat.

With perfectly inelastic balls we see from equation (5.3) that the final kinetic energy is $\frac{(mv_1 + Mu_1)^2}{2(m + M)}$. Hence, if the initial momenta are equal and opposite (*i.e.* $Mu_1 = -mv_1$), there is no final kinetic energy, and the balls remain at rest at the point of impact. In any case they remain together, for equations (5.3) show that when e vanishes, $u_2 = v_2$.

The subject of impact is obviously of practical importance in connection with collisions of vehicles and other bodies.

SOLUTION

Solution is the interpenetration of two bodies, their molecules being thoroughly intermingled. In general, this may happen whether the bodies are solid, liquid, or gaseous, but the term "solution" is often restricted to the case of the interpenetration of a liquid by a solid or gas. Two solids, however, when kept in contact with one another, may slowly interpenetrate, and one may even pass right through the other. An alloy is a partial example of such interpenetration, but this is generally formed when the components are in the liquid

state. Again, certain solids are able to take up large volumes of gases—a phenomenon known as *occlusion*. Palladium, for instance, occludes hydrogen strongly, and certain forms of charcoal can occlude most gases to a surprisingly large degree. This phenomenon is often used for the purpose of obtaining high vacua. The charcoal, in small fragments, is cooled by liquid air, for its power of occlusion rises rapidly with lowering of temperature, and in these circumstances almost all gases except the rare gases (and these also to a slight extent) enter it, no matter how small their pressure. By this means a space can be almost completely evacuated.

Adsorption

With most solids and gases, however, interaction is mainly a surface phenomenon. Usually, when a gas is contained in a solid vessel, the surface of the vessel becomes coated with a layer of the gas which is not detachable without great difficulty. This phenomenon is known as *adsorption*. The layer is extremely thin—usually only one molecule thick—and the amount of gas contained in it varies with the material of the vessel and the gas, with the partial pressure of the gas, and with the temperature. Adsorption is a source of much trouble when high vacua or gases of high purity are required, since a change of physical conditions may bring some of the adsorbed gas into the vessel. Prolonged heating, often to a high temperature, is necessary before the layer can be completely removed.

Liquid Mixture

Interpenetrability of liquids and liquids is usually called *mixture*. Most liquids can mix with one another (they are then said to be “miscible in all proportions”), but not all. Thus we are familiar with the appearance of oil floating on water. Some pairs of liquids are partly miscible. For example, if ether and water are poured into the same vessel there is partial mixing, but there will be an excess of one of

the liquids which will remain unaffected. We may say that water will dissolve ether, or ether water, to a limited degree.

Solution in Liquids

The commonest form of solution, however, is that of solids and gases in liquids. It is probable that liquids will dissolve any solid or gas to some extent, though often the amount is so small that the solid or gas may be said to be insoluble. The dissolving liquid is called the *solvent*, and the substance dissolved the *solute*.

The amount of a substance which can be dissolved by a particular amount of a particular solvent is limited. When the maximum amount has been dissolved the solution is said to be *saturated*. The quantity required for saturation varies with the nature of the solute and of the solvent, and in general it varies also with temperature. The process of solution is speeded up by increase of temperature, and the amount of solute required for saturation is usually increased thereby. An important exception is the solution of common salt in water. Here the amount which can be dissolved is almost independent of the temperature.

Supersaturation

It is possible to obtain a "supersaturated" solution of certain substances (e.g. sodium sulphate) by first obtaining a saturated solution in hot liquid and then carefully cooling it. Normally, the excess of solute (supposing, as in the usual case, that the amount required for saturation increases with rise of temperature) then crystallizes out, but with care it may be kept in solution. The introduction of even an extremely minute fragment of the solute, however, or even mechanical shaking, suffices to precipitate solute until the normal saturation strength is reached.

Detailed discussion of the phenomena of solution is beyond the range of this book. It appears that in dilute solutions the dissolved substance behaves in many respects like a gas, and

indeed the distances between its molecules are comparable with those proper to the gaseous state. A very significant classification of solutions is the division into conductors and non-conductors of electricity. It is believed that in the former the molecules of the solute are dissociated into their constituent atoms, and no longer exist whole as in their original solid state.

OSMOSIS

A very important phenomenon—important in physiology as well as in physics—is that known as *Osmosis*. It is closely connected with the general theory of solution. The fundamental fact of osmosis is that certain membranes are permeable by some liquids and not by others ; in particular, they may be permeable by a pure liquid but not by a solution, or by solutions of some substances but not by those of others. In such cases, if the membrane has on one side a liquid to which it is permeable, and on the other a liquid to which it is impermeable, the former will pass through the membrane until a definite excess of pressure is set up on the other side. This excess of pressure is known as *osmotic pressure*.

This phenomenon was discovered nearly two hundred years ago, when it was found that a bladder of alcohol immersed in water swelled considerably, while a bladder of water immersed in alcohol shrank. It may be observed very simply by fixing a sheet of parchment over the lower end of a tube into which sugar solution is poured. On immersing the vessel in water, keeping the upper portion of the tube above the surface, the level of the solution in the tube may be seen to rise.

Semi-Permeable Membranes

Membranes which are permeable to water but not to most aqueous solutions are called *semi-permeable* membranes. A convenient form of semi-permeable membrane is obtained by placing a porous earthenware pot, containing a dilute solution

of copper sulphate, in a solution of potassium ferrocyanide. Both solutions pass into the pores of the pot, and, where they meet, a film of the gelatinous ferrocyanide of copper is formed, which is a semi-permeable membrane. The function of the pot is merely to protect the membrane, which easily collapses when unsupported : the pot itself is permeable by most liquids.

Crystalloids and Colloids

There are other membranes which are permeable not only by water but also by certain solutions, though not by all : parchment paper is a good example. This fact allows us to make a very important classification of solutions into *crystalloids* and *colloids*, crystalloids being those substances whose solutions will pass through such membranes, and colloids those whose solutions will not. Crystalloids include most salts which form definite crystals, while colloids include amorphous, gelatinous substances. A colloidal solution appears to be not a complete interpenetration of the molecules of solvent and solute, but a suspension of exceedingly small globules of the solute in the solvent. A mixture of crystalloids and colloids can be completely separated by means of a membrane of such a substance as parchment paper.

Osmotic Pressure

For a given solution the osmotic pressure has a definite value, which may be found by means of the apparatus illustrated in Fig. 26. Here A is a porous pot containing the solution, which is placed in an outer vessel of pure water. Water enters the solution until the liquid in the capillary tube rising from A reaches a definite level. The height of this above the level of pure water measures the excess pressure in the pot, *i.e.* the osmotic pressure.

By such means it is found that, for dilute solutions, the osmotic pressure is proportional to the strength of the solution (*i.e.* to the density of the solute in the solution) and also to the absolute temperature (see Chap. VII.). This is an example

of the fact mentioned above, that the solute in a dilute solution behaves like a gas ; for the pressure of a gas also is proportional to the density and the absolute temperature (see Chap.VII.). Moreover, the constant of proportionality is found to be the same in the two cases.

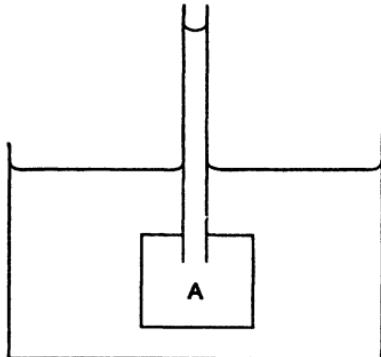


FIG. 26

Measurement of Osmotic Pressure

water to the solution. But more molecules make the latter passage, because more water molecules strike the outer wall of the pot. This is because many of the molecular impacts on the inner wall are those of solute molecules which cannot get through. Consequently water will pass from the outside to the inside until the number of impacts per second of *water* molecules is the same on both sides. The impact of the solute molecules on the inside then constitutes the excess (osmotic) pressure. Clearly, the stronger the solution, the greater will be the osmotic pressure.

DIFFUSION

Diffusion of Liquids

If two liquids which are miscible in all proportions are carefully placed in contact with one another without the interposition of a membrane, they will slowly mix until ultimately they form one homogeneous liquid. The tendency

of gravity to keep the dense liquid at the bottom merely delays the mixing without preventing it, though it may delay the process by several years. This process is known as *diffusion*. It may easily be observed by pouring water very carefully on a strongly coloured solution. The progress of the diffusion is indicated by the subsequent change of colour at different parts of the column.

The law of diffusion was first stated by Fick: it is as follows. Fig. 27 represents a cylindrical vessel containing a strong solution of a salt at the bottom, and pure solvent or weak solution at the top, the strength of the solution in the intervening liquid gradually increasing from the top downwards. Diffusion of salt will take place from below upwards. Fick's law states that the amount which passes a square centimetre at any level is proportional to the gradient of concentration at that level; *i.e.* if the strength of the solution is n gm. of salt per cc. at a level x cm. from the top, the rate of diffusion is $k \frac{dn}{dx}$, where k is a constant known as

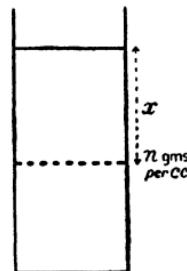


FIG. 27

Fick's Law of Diffusion

the *diffusivity* of the solution. It can be shown to follow from this that the time required for diffusion through any distance is proportional to the square of that distance.

Among ordinary salts the diffusivity varies over a range of three- or four-fold. Common salt has a good average value: a strong solution of this substance at the bottom of a column of water 1 metre high would diffuse uniformly throughout the liquid in about twenty-one years. With a column one-fifth of this height, however, the time would be less than one year, an account of the relation just stated.

Diffusion of Gases

The diffusion of gases takes place according to the same law as the diffusion of liquids, but very much more quickly.

With gases the concentration n is represented by the density ρ , and we have therefore the expression $k \frac{d\rho}{dx}$ for the rate at which either of the two gases concerned crosses any level. k in this case is called the *interdiffusivity*, for the gas here diffuses into another gas, and not into a solvent.

The difference in speed of diffusion in gases and liquids may be indicated by comparing the rate already given for sodium chloride solution with the following result obtained by Graham for the diffusion of CO_2 into air. CO_2 was placed at the bottom of a cylinder 57 cm. high, so that it occupied one-tenth of the height, and the remainder was filled with air. Mixing was complete after about two hours.

Explanation of Diffusion

On the molecular theory of matter, these results are quite intelligible. The molecules of a gas are much freer to move than those of a liquid, though the latter have some freedom. Consequently a certain proportion will wander into the space around. This happens, of course, whether there is another gas or liquid there or not, and one result of diffusion of liquids is that every liquid has around or above it a certain quantity of its own vapour. Detailed consideration shows that the rate of diffusion of a gas should depend on the density of the gas in the manner found by experiment.

The rate of interdiffusion is important in relation to the rate of evaporation of a liquid. The air above a water surface, for example, can hold only a certain amount of water vapour, known as the saturation amount (see Chap. XI.), and when this value is reached evaporation stops. In ordinary circumstances, however, such a condition is never permanently maintained, for the vapour diffuses into the surrounding air, thus making it possible for more water to evaporate, so that the rate at which a lake, for example, will dry up in a period of drought depends on the rate of diffusion of water vapour into the air.

Diffusivity Coefficient

The coefficient k , with both liquids and gases, is not exactly constant, but varies somewhat with the composition of the mixture. It depends also on temperature—increasing, as we should expect, with rise of temperature—and, in the case of gases, is inversely proportional to the pressure of the mixture.

The variation of k with different gases affords a useful means of separating the gases in a mixture. For this purpose it is convenient to use a porous vessel through which the gases can gradually percolate. It is found that each makes its way through as though the other were not there, so that the one with the highest value of k is transmitted most rapidly. The rate of passage is directly proportional to the difference of pressure on the two sides of the pot, and equal volumes of different gases pass through in times proportional to the square roots of their molecular weights. This result also is quite in accordance with the molecular theory of gases.

Thermal Effusion

A similar phenomenon is that known as *thermal effusion*. If two gases are at the same pressure on opposite sides of a porous partition, but one is hotter than the other, then gas will pass from the colder to the hotter side. The reason for this result, which is surprising at first sight, is that since the pressures are equal, the hotter gas must have the lower density (see p. 126), and the direction of diffusion is that of the density gradient.

VISSCOSITY

When a fluid moves in obedience to an external force, it does not do so as a rigid body would, but one layer retards the motion of the next, in a manner somewhat akin to that of the frictional resistance of solids. This phenomenon is known as *viscosity*. It occurs with both liquids and gases, but since the results of the resistance are different in the two cases, it is convenient to treat them separately.

Viscosity of Liquids

The viscosity of liquids can be understood best by taking a particular case—that of a liquid flowing along a tube. To make the matter definite, suppose the tube is horizontal and the liquid is forced along it by a constant pressure at one end—created, say, by a constant head of the liquid, as in Fig. 28. In these circumstances the flow may be either

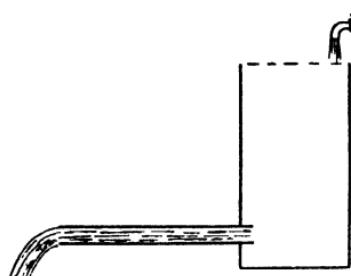


FIG. 28

Steady flow of liquid along tube

of two kinds—steady and turbulent—and which will occur depends on the velocity of flow, which, in its turn, depends on the head of liquid. There is a certain velocity, known as the *critical velocity*, which is such that a liquid whose velocity exceeds it does not flow smoothly but forms eddies and moves very irregularly. Such motion is said to be *turbulent*. If the ve-

locity is less than the critical velocity, however, the motion is *steady*, each small element of the liquid moving parallel to the axis of the tube throughout the flow. This may be shown very well by introducing a small fragment of colouring matter, so that the lines of flow are clearly exhibited. The critical velocity is found to be inversely proportional to the radius of the tube, so that in narrow tubes we can have smooth motion with high velocities.

We shall consider only the case of steady flow. It is found that the velocity of an element of the liquid is not constant at all parts of the tube, but increases as we go from the walls to the axis. There is, in fact, strong reason to believe that the layer of liquid actually in contact with the walls is stationary. This difference of velocity indicates a tangential force exerted by one layer of liquid on the next ; the faster liquid tends to hasten the slower, and the slower

tends to retard the faster. The flowing liquid may be regarded as made up of a series of concentric cylindrical liquid tubes of small thickness and gradually increasing radius, a cross-section of which is shown in Fig. 29, where the outer thick ring represents the glass tube. If these component tubes be numbered from 1 to n , starting with the outermost, then tube 1 (if the thickness is sufficiently small) remains at rest, while the central core, n , has the greatest velocity of all. Tube 3, say, tends to pull the slower tube 2 with it, and to drag back the faster tube 4.

Coefficient of Viscosity

There is evidence that the magnitude of this viscous force is proportional to the rate at which the velocity changes as we go from the walls to the axis of the tube. Suppose that the central layers of two consecutive concentric tubes are unit distance apart, and that there is unit difference of velocity between them. Then the tangential force on each unit of area at the surface where they meet is called the *coefficient of viscosity*, or, more briefly, the *viscosity* of the liquid. If this be denoted by η , then it follows that the force over an area

A at a place where the velocity gradient is $\frac{dv}{dr}$, is

$$F = \eta A \frac{dv}{dr} \dots \dots \dots \dots \quad (5.4)$$

This tangential force has clearly something of the nature of a shearing stress. We have said (p. 52) that liquids have negligible rigidity, and that is true in the sense that they will change shape in obedience to the smallest force. On the other hand, however, they always do so under protest, so to speak, and when moving under the influence of such a force they resist its operation. This resistance is their



FIG. 29

Imaginary section of tube containing flowing liquid

viscosity, and while we cannot measure it as an ordinary modulus of elasticity, by the ratio of the stress to the consequent strain, we can do so by the ratio of the stress to the consequent velocity gradient.

Measurement of Viscosity

It is possible to determine the coefficient of viscosity of a liquid by observing its steady flow along a horizontal tube. It can be shown that, if the head of liquid be kept constant, the volume flowing in unit time is given by the expression

$$Q = \frac{(P_1 - P_2)\pi r^4}{8\eta l} \quad \dots \quad \dots \quad \dots \quad (5.5)$$

where P_1 and P_2 are respectively the pressures at the entrance and exit ends of the tube, and r and l are the radius and length of the tube. This is known as *Poiseuille's formula*.

One point might be mentioned about this formula. It appears that if $\eta = 0$, the volume flowing is infinite. The explanation of this apparent absurdity is that the equation applies only to *steady* flow, which is flow with constant velocity, and if there were no viscous resistance to the difference of pressure, such a state would never be attained ; the motion would be accelerated. Hence we cannot properly extrapolate to the case of a non-viscous liquid. This, of course, is not a practical limitation, since such a liquid does not exist.

Explanation of Viscosity

From the point of view of the molecular theory, viscosity may be looked upon somewhat as follows. Molecular forces keep stationary the layer of liquid in contact with the tube, while the next layer is free to move. There is, however, a constant interchange of molecules between adjacent layers, because the molecules of a liquid have a certain freedom of movement. Hence the free layer is constantly receiving

molecules from a layer which, on the whole, has no movement in the direction of flow, so that the average velocity of flow of its molecules is reduced. Interchange between this and the next layer similarly reduces the average velocity of that layer, and so on. From this point of view the drag which each layer exerts on its neighbour is due to the exchange of slow molecules for faster ones.

Motion of Solids through Liquids

An important example of the effect of viscosity is provided by the motion of solids through liquids. If we have a solid body falling through a liquid, the layer of liquid in contact with the solid shares its motion, just as the liquid flowing along a tube has the layer next to the tube relatively stationary. The distant liquid, however, is not moving, so that we have a velocity gradient set up, and therefore a viscous resistance to the motion. It can be shown that this ultimately brings the velocity of the solid to a constant value. It was proved by Stokes that a sphere, of radius a , moving with a small velocity v through a fluid of viscosity η , experiences a resistance on this account equal to $6\pi\eta av$. It follows from this that if the velocity increases, the resistance increases also, and when it equals the force which moves the body (gravity in the ordinary case of falling bodies) there will be no resultant force, and therefore no acceleration.

Suppose the sphere is falling under its own weight, which is $\frac{4}{3}\pi a^3 \rho g$, where ρ is its density. Then, if the liquid through which it moves has density σ , the weight is effectively $\frac{4}{3}\pi a^3 g(\rho - \sigma)$, and by equating this to $6\pi\eta av$ we see that the terminal velocity, v , is equal to $\frac{2}{9} \frac{a^2 g}{\eta} (\rho - \sigma)$. This equation applies to gases as well as liquids : it can be used, for instance, to determine the terminal velocity of a small raindrop falling through air, or, alternatively, the mass or density of particles falling through fluids if the velocity be observed. In all such applications, however, care must be taken to see that

the velocity never exceeds the limiting value for which Stokes's formula holds.

As a means of determining the masses of very small or light particles, this method is limited by the weakness of the Earth's gravitational field. For given small values of a and $(\rho - \sigma)$, the terminal velocity is proportional to g , and in many problems of interest, in which colloidal particles or even single organic molecules of high molecular weight are concerned, it is too small for practical purposes. This difficulty has been overcome very ingeniously in an instrument known as the *ultracentrifuge*, developed mainly by Svedberg, in which a strong field of centrifugal force is substituted for the gravitational field of the Earth. A liquid containing the particles in suspension is made to rotate very rapidly, the particles thus being subjected to a centrifugal acceleration, $\omega^2 r$, where ω is the angular velocity and r the distance of a particle from the axis of rotation. This must be substituted for g in the above formula, and in modern instruments values of the order of a million times g are attained. The particles move along the radius of the rotating liquid at rates depending on their mass and density. In this way the molecular weights of certain proteins have been determined, and the method is useful also as a means of separating particles of different mass.

Viscosity of Gases

Gases possess viscosity for the same reason as liquids, and the coefficient of viscosity is defined similarly. In terms of the molecular theory, we have the same interchange of molecules between the slow and faster moving portions, and so the same kind of viscous drag. Poiseuille's formula, however, does not hold without modification for a gas because, although the gas experiences the same force as a liquid, its behaviour under the action of the force is different. A liquid being effectively incompressible, the condition of steady flow (namely, that the same mass passes each cross-section of the

tube per second, so that there is no accumulation or deficit of liquid anywhere in the tube) is equivalent to the condition that the same *volume* passes each cross-section per second. With a gas, however, which is very easily compressible, the same mass has a different volume at the entrance end, where the pressure is higher, from that at the exit end. The result of introducing this modification is that we obtain the following formula for gases :

$$P_1 V_1 = P_2 V_2 = \frac{(P_1^2 - P_2^2)}{16 l\eta} \pi r^4 \dots \dots \quad (5.6)$$

where V_1 and V_2 are respectively the volumes entering and leaving the tube per second, and the other symbols have the same meanings as before.

Viscosity and Pressure of a Gas

A very remarkable fact about the viscosity of a gas is that it is practically independent of pressure over a wide range. This may be understood if we think of the matter from the point of view of the molecular theory. We have seen that viscosity arises from the interchange of fast and slow molecules across the boundary between two layers of the gas. Let us suppose the pressure is now diminished (without, of course, any change of temperature). The volume of the gas must then increase, and the distances between the molecules must therefore increase also. In that case we should expect fewer molecules to cross each unit area of the boundary. That would be so, but for the fact that the decreased density allows molecules from a greater distance to reach the boundary, the attenuation of the gas giving them a larger path of free movement without collision. These two effects just neutralize one another, with the result that the same interchange of molecules takes place as before and the viscosity is therefore the same.

This ceases to hold when the average distance between the molecules is so great as to be comparable with the dimen-

sions of the containing vessel. If the pressure is lowered below this point the viscosity decreases. But from atmospheric pressure down to a few millimetres of mercury there is no appreciable variation. It follows, for example, that if we have a pendulum swinging in air, we do not diminish the viscous resistance to its motion by decreasing the pressure until we get down to very low pressures.

Viscosity and Temperature

The viscosities of gases and liquids are affected oppositely by change of temperature. Increase of temperature diminishes the viscosity of liquids (the effect of heat on syrup is a homely example of this) but increases that of gases. This may be explained in the following way. Increase of temperature corresponds to an increase of kinetic energy of the molecules and so facilitates the interchange of molecules which gives rise to the viscous drag. With gases this is the only effect, and so the viscosity of gases rises with temperature. With liquids, however, this effect is overpowered by the diminution of cohesion of the liquid. The forces of attraction between the molecules are diminished, so that one layer can slip more easily over the next.

EXERCISES

1. Describe the movements of two spheres after a head-on collision when the coefficient of restitution is (a) 1; (b) $\frac{1}{2}$; (c) 0. If the mass and initial velocity of one are respectively 20 gm. and 30 cm. per sec., and those of the other 40 gm. and -20 cm. per sec., what will be the final velocities in the three cases, and how much kinetic energy, if any, will be lost by the impact?
2. Describe the phenomena of osmosis, and explain why liquid can pass through a membrane in opposition to the pressure gradient.

3. State Fick's law of diffusion of liquids. If a liquid takes t days to diffuse uniformly through a distance l cm., how long will it take to diffuse through $2l$ cm.?
4. Give an explanation of the viscosity of liquids on the molecular theory, and state Poiseuille's equation for the flow of liquid along a horizontal tube. Why is the equation different for the flow of a gas? If 150 c.c. of liquid flow per second through a tube 30 cm. long and 0.5 cm. in diameter, when the pressure difference at the ends is that of a column of water 30 cm. high, what is the viscosity of the liquid?

PART III
HEAT

CHAPTER VI

HEAT AND TEMPERATURE

THE NATURE OF HEAT

The subject of heat originates in our sensations of “hot and cold.” These, however, are not directly measurable, so we try to account for them by assuming that there is something which we call *heat* which passes from a hot body to a colder one when they are placed in contact with or near one another.

This “heat” is not, of course, directly perceptible : we experience only the effects of receiving or parting with it. Our ideas of its character must therefore be derived by inference from the behaviour of bodies. The early workers in the subject had thus to make experiments without any very clear idea of what was occurring in the bodies they were examining, and to assist their thoughts they imagined that heat was an invisible material fluid, which they called *caloric*, which tended to pass from a hot to a cold body.

This idea has much to recommend it, and it enabled some important advances to be made, but it broke down in the face of certain facts. One of the most important of these was the fact that heat could be developed by friction ; *i.e.* by rubbing two bodies together. If only a small amount of heat could be developed in this way it would be possible to say that the capacity of a body for heat is diminished by rubbing, so that some heat escapes. We can, however, apparently produce an indefinitely large amount of heat by friction ; so long as we continue rubbing, heat continues to

be produced. We can scarcely suppose that bodies can contain an infinite amount of a material fluid, and the idea of caloric had therefore to be abandoned.

Heat and Motion

The modern idea is suggested directly by this experiment. Since motion (the motion of rubbing) produces heat, we may suppose heat to be a form of motion. There is then no difficulty in understanding that the more we rub, the more heat is produced. The motion that constitutes heat is, of course, motion of the atoms or molecules of a body: the hotter the body, the faster the molecules move.

This conception of molecular motions is of great help in understanding the phenomena of heat, or *thermal* phenomena, as they are called. Nevertheless, we must remember that a great deal of pioneer work in the subject was done without its assistance, and this fact may answer some otherwise puzzling questions. For instance, if heat is the energy of motion of molecules, it would be natural to measure heat in terms of this energy; unit quantity of heat would be unit quantity of molecular energy. This is not done, however: unit quantity of heat, as we shall see, is defined in terms of the amount by which a particular substance—water—gets hotter when it receives heat. This somewhat arbitrary measure is now too firmly established to be discarded, and we must accept it. It has at least the advantage that it can be directly determined, whereas the energy of molecular motions can be measured only indirectly.

Temperature

When we say that one body is hotter than another, we mean that heat will flow from the former to the latter. We do not, however, mean that the former contains more heat than the latter; it may contain less. The state of being hotter or colder than a given standard we call the *temperature* of a body, and we then say that heat flows from a body of higher to a

body of lower temperature, no matter how much heat the bodies may contain.

A partial analogue of the relations between heat and temperature is provided by the relations between a quantity of water and the height of the water above the ground. Water flows from a higher to a lower level, irrespective of how much water is already at the two levels, and, if it meets with no obstacle, continues to do so until all the water is at the same level. Water will no more flow uphill than heat will flow against a temperature gradient.

The matter may be understood quite well in terms of our conception of heat as molecular energy. The heat which a body contains is the total energy of its molecules : the temperature of the body is proportional to the *average* energy of one of its molecules. Suppose two bodies—for simplicity, of the same material, so that all the molecules have the same mass—are brought into contact. If they are at the same temperature, the more massive will have the greater amount of heat, because it has more molecules. If, however, the less massive is at a higher temperature, its molecules will be moving faster, and on striking the molecules of the other body, they will impart some of their energy thereto. In this way the average energy (*i.e.* the temperature) of the more massive body will be increased, and that of the less massive body will be decreased. Heat will thus pass from the body whose molecules are moving faster to the body whose molecules are moving slower, although the latter may already contain the greater total energy (*i.e.* the greater heat).

Measurement of Temperature

We cannot measure the average kinetic energy of molecules directly, so we have to choose another way of measuring temperature. We do so by looking for some easily measurable *effect* of change of temperature, and then measure the temperature change by the change in this effect. For example, it is

found that (with a few exceptions) when the temperature of a body rises the body expands, and when the temperature falls the body contracts. We might therefore obtain a measure of temperature by taking equal amounts of expansion of a selected body to indicate equal increments of temperature.

Again, a rise of temperature causes the electrical resistance of a body to increase, and we might obtain another measure of temperature by taking equal increments of resistance of a selected body to indicate equal increments of temperature. Any other measurable effect of temperature might similarly be taken as a basis of measurement of temperature. It does not follow, of course, that equal changes of temperature on any of the scales so obtained will correspond to equal changes of the average kinetic energy of the molecules of a body, but here again we adopt our arbitrary scale for ordinary purposes, and only deduce the molecular energy when for special reasons it is necessary to do so.

Mercury-in-glass Thermometer: The commonest temperature scale is that provided by expansion, and the commonest instrument recording this scale is the *mercury-in-glass thermometer*. This, as is well known, consists of a bulb and a capillary stem containing mercury. As the thermometer becomes hotter the mercury and the glass both expand, but the former expands the more, and so rises in the stem. The amount of expansion, indicated by a scale engraved on the stem, measures the temperature.

The unit of temperature increase is called the *degree*; it is obtained in the following way. The thermometer is placed in a mixture of distilled water and ice, which mixture is found by experience always to have the same temperature, and a mark is made at the level of the mercury in the stem. The instrument is then transferred to the steam issuing from distilled water boiling at normal atmospheric pressure (760 mm. of mercury), and when the mercury comes to rest another mark is made at its new level; this, of course, will be higher

than the former one. These two marks correspond to two definite temperatures (they are known as the *fixed points* of the thermometer), and it only remains to assign numbers to them in order to obtain our scale of temperature.

This is done in three ways, giving three different scales. On the *Centigrade* scale, which is generally used for scientific purposes, the temperature of melting ice is called 0 degrees (written 0°), and that of steam is called 100° . The interval between the marks on the stem is divided into a hundred equal parts, and each is called a degree. On the *Fahrenheit* scale, used in clinical work and for domestic purposes in Britain, the fixed points are given the numbers 32 and 212 respectively, and the interval is divided into 180 equal parts, or degrees. The *Fahrenheit* degree is thus a smaller increment of temperature than the *Centigrade* degree. On the *Réaumur* scale, used in Germany, the fixed points are called 0° and 80° , and the *Réaumur* degree is therefore greater than the *Centigrade* degree. On each instrument the scale is extended below the temperature of melting ice and above that of steam by continuing the graduations of the stem uniformly beyond the fixed points.

Temperatures on these scales are indicated by the appropriate number followed by the sign $^{\circ}$ C., $^{\circ}$ F., $^{\circ}$ R., as the case may be. From the way in which the scales are constructed it is easy to see that if the temperature of a particular body is given as c° C., f° F., r° R. respectively by the three thermometers, c , f , and r must satisfy the following relation:

$$\frac{c}{5} = \frac{f - 32}{9} = \frac{r}{4} \dots \dots \dots \quad (6.1)$$

Gas Thermometers: For some purposes a gas is used instead of mercury as a thermometric substance. Gases expand much more than liquids for the same rise in temperature, and therefore provide much more sensitive instruments. They are usually graduated according to the *Centigrade*

scale, and their readings do not differ greatly from those of the mercury-in-glass thermometer. They have the great advantage, however, that they can be used over a much wider range. Mercury freezes at -39° C., and boils at 358° C., so that it is very restricted in its application. Some gases, however, remain gaseous at extremely low temperatures, and there is no upper limit to their suitability as thermometric substances. The unwieldiness of gas thermometers is their chief defect. For this reason, when a portable instrument is required to measure temperatures slightly below the freezing-point of mercury (*e.g.* in the Arctic regions), alcohol-in-glass thermometers are usually preferred.

Use of Thermometers: The method of using these thermometers is the same for all. They are immersed as completely as possible in the substance whose temperature is required, and the temperature is then indicated by the reading on the graduated stem. Correction may have to be applied for the portion of the stem which remains exposed if great accuracy is desired.

Modes of Transfer of Heat

We have spoken of the transfer of heat from a body of higher to one of lower temperature when they are brought into contact. This process is called *conduction* of heat. The faster molecules of the hot body collide with the slower ones of the other, and impart some of their energy thereto, with the result that they themselves are somewhat slowed down. The temperatures therefore tend to approach equality.

This, however, is not the only way in which a hot body can transfer heat to a colder one. It does so spontaneously, even if the bodies are separated by empty space. An obvious example of this is the fact that the Sun warms us over a distance of 93 million miles, the space between being almost empty of matter. This process is called *radiation*, and is attributed to the setting up of waves in a medium called

the *ether*, which is conceived to exist everywhere, occupying even the spaces between the molecules of bodies. The vibrations of the molecules originate these waves, which then travel through the ether at a rate of about 186,000 miles (3×10^{10} cm.) a second, and on meeting another body give up their energy to the molecules of that body, which therefore acquire heat.

This conception of waves in the ether serves to explain many of the facts of radiation of heat, but the student is warned against thinking that it represents in detail what actually occurs. The subject is a very difficult one, and while we shall accept the hypothesis of ether waves in order to give an intelligible account of the phenomena of radiant heat and of light also, it is necessary to state that it is not now believed to give a final account of the matter. We must be careful not to apply it beyond the range of phenomena which it explains satisfactorily. It is not always necessary in science to think in terms of what is believed to be the conception nearest to the truth of the matter. Just as we find it convenient to speak of sunrise, although we believe the phenomenon indicated to be more accurately expressible in terms of the rotation of the Earth, so we often exercise our right to express phenomena in terms of an easily picturable imaginary process so long as it does not lead us astray. We shall meet with other examples of this practice.

There is a third mode of transfer of heat known as *convection*, which is possible only with fluids in a gravitational field. If a liquid is heated from below, the lower parts expand and so become less dense. They therefore float upwards and the cooler liquid above descends to take their place. This in its turn is heated and rises, and so a general circulation takes place, the heat being quickly dispersed throughout the body of the liquid in this way. If the liquid is heated at the top, convection cannot occur, and the much slower process of conduction is the only way by which heat can spread throughout the liquid.

EXERCISES

1. Explain the difference between heat and temperature. How are the two ideas interpreted on the molecular theory of matter?
2. Give an account of the various ways in which a scale of temperature measurement may be constructed, and discuss their relative advantages.
3. Calculate the Centigrade temperatures corresponding to 98.4° F. and 0° F., and the Fahrenheit temperatures corresponding to 50° C. and -32° C.
4. What temperature is represented by the same number on the Fahrenheit and Centigrade scales, and for what temperature does the Fahrenheit thermometer read twice as high as the Centigrade thermometer?

CHAPTER VII

THERMAL EXPANSION

EXPANSION OF SOLIDS

HAVING decided how to measure temperature, we can now study the expansion of bodies with rise of temperature. The general principles are the same for all bodies, but the differences in detail between solids, liquids, and gases are so important that the three states of matter must be considered separately.

Effects of Expansion

The expansion of solids has many important effects in nature. Most bodies are exposed to the alternations of heat and cold resulting from the succession of day and night and of the seasons, and it is sometimes necessary to take special precautions against the resulting expansion. For instance, in laying railway lines, gaps have to be left here and there so that in hot weather the lines shall not buckle: it has been calculated that the natural changes of temperature might cause a change of length of over 200 yards in the 400 miles or so of line between London and Edinburgh, and, unless this were allowed for, accidents would result. Again, we are all familiar with the breaking of thick glass vessels when hot water is poured into them. The inside is heated and expands, while owing to the poor conductivity of glass for heat, the outside remains cool. A strain is thus set up, which breaks the vessel. The effect of expansion on delicate measuring instruments must always be taken into account. For example,

clock pendulums and the balance wheels of watches have to be specially constructed so that change of temperature shall not impair their accuracy.

The force of thermal expansion and contraction is very great, and has been used to break down heavy structures. If two walls, for example, are connected by a hot metal rod, contraction of the rod on cooling exerts sufficient force to break them down.

Linear Expansion

A solid when heated expands in all directions, but we are often (as in the case of a measuring rod, for example) interested only in the expansion in one direction. This is called the *linear expansion* of the solid.

It is found by experiment that when a rod is heated, its linear expansion is proportional to the initial length of the rod and to the rise in temperature. Thus, if we take the length at 0° C. as a standard length, and call it L_0 cm., then, when the rod is heated to t° C., it expands by aL_0t cm., where a is a constant for the rod. For another rod, made of a different material, the same relation is true, but a has a different value : it is constant only for a given material. a is called the *coefficient of linear expansion* of the material, and is a definite property of a substance, like its atomic weight or melting point. The coefficient of linear expansion is clearly the length by which a rod 1 cm. in length at 0° C. expands when heated to 1° C.

If we denote the length of the rod at t° C. by L_t , then obviously

$$L_t = L_0 + L_0 at = L_0 (1 + at). . . . (7.1)$$

This is a fundamental equation connecting the length of the rod at any temperature with its length at 0° C. and with the coefficient of linear expansion.

Measurement of Linear Expansion : The methods of measuring a are quite straightforward. The length of a rod of the

substance is measured, usually by reading microscopes, when the rod is in melting ice and afterwards when it is in boiling water. The values obtained are L_t and L_0 , respectively. t is 100, so that equation (7.1) can be used to determine α .

The values of α are very small for all solids, and thermal expansion of solids is therefore not normally of great importance unless large changes of temperature or rods of great length are in question. The following are a few typical values :

Copper	0.0000168
Steel	0.0000110
Brass	0.0000187
Glass	0.0000083
Platinum	0.0000089

Cubical Expansion

If the expansion of a solid in all directions is of interest we must consider the change of *volume* and not merely the change of length. In that case we have a *coefficient of cubical expansion*, as it is called, which is defined as the volume by which a solid, 1 c.c. in volume at 0° C., expands when heated to 1° C. By the same reasoning as before, we obtain the equation

$$V_t = V_0 (1 + \beta t) \quad \dots \quad \dots \quad \dots \quad (7.2)$$

where β is the coefficient of cubical expansion, and V_t and V_0 are the volumes at t° C. and 0° C. respectively.

It can be proved that β is approximately equal to 3α for any solid whose coefficient of linear expansion is the same in all directions. (Non-crystalline substances satisfy this condition, but crystals often have different coefficients for different directions). For, let us imagine a rectangular block of a solid whose sides at 0° C. are of length a , b , and c cm. respectively. Then $V_0 = abc$ cc. On being heated to t° C.,

the sides expand to $a(1+at)$, $b(1+at)$ and $c(1+at)$ cm., respectively, so that

$$V_t = abc(1+at)^3 = V_0(1+at)^3. \quad \dots \quad (7.3)$$

Comparing this with (7.2) we obtain

$$1 + \beta t = (1+at)^3 = 1 + 3at + 3a^2t^2 + a^3t^3. \quad \dots \quad (7.4)$$

Hence

$$\beta = 3a + 3a^2t + a^3t^2.$$

Now unless t is very large, the last two terms on the right-hand side are very small compared with the first term, and may be neglected. We then have

$$\beta = 3a. \quad \dots \quad \dots \quad \dots \quad \dots \quad (7.5)$$

This relation enables us to deduce the coefficient of cubical expansion from measurements of the coefficient of linear expansion.

Variation of a

If we require great accuracy, or are concerned with fairly large changes of temperature, we must allow for the fact that a is not quite constant, but varies as the temperature rises: that is to say, the expansion of the rod when heated from 0° C. to 1° C. is not *exactly* the same as the expansion when it is heated from 100° C. to 101° C., say. It is sometimes necessary, therefore, to use, instead of equation (7.1), the more exact equation

$$L_t = L_0(1+at+a't^2) \quad \dots \quad \dots \quad \dots \quad (7.6)$$

where a' is another constant, much smaller than a . This equation may also be written

$$L_t = L_0[1 + (a + a't)t] \quad \dots \quad \dots \quad \dots \quad (7.7)$$

in which we see that the expression $(a + a't)$ —a function of the temperature—plays the part of a in the simpler equation. We may therefore say that the coefficient of linear expansion

is a function of the temperature which is almost constant except at high temperatures. A corresponding result holds, of course, also for β .

Expansion of Alloys

The expansion of an alloy cannot be easily deduced from the expansions of the component metals. Apparently two influences come into play. Thermal expansion in general may be attributed to the greater amplitude of the molecular motions as they acquire heat energy, but accompanying this effect there may be a greater possibility of compact arrangement of the molecules, just as shaking makes a powder settle into a smaller volume in the containing vessel. When these two effects are equal and opposite, the size of an alloy may be unaffected by heat. This possibility is practically realized in a nickel-steel alloy known as *invar*, a bar of which remains effectively constant in length over a considerable range of temperature. This is an extremely valuable property, of which much use is made in the construction of instruments of precision.

Among optical materials quartz has a much smaller coefficient of expansion than glass. It is therefore used when the effects of expansion are likely to be seriously detrimental.

EXPANSION OF LIQUIDS

Real and Apparent Expansion

In dealing with liquids we have no interest in linear expansion, and consider cubical expansion only. Equation (7.2) holds good for liquids as well as solids, but we have the complication here that a liquid must always be contained in a vessel, and when it is heated the vessel of course becomes heated also and itself expands. Liquids, however, have higher coefficients than solids, so that—as in a mercury-in-glass thermometer, for instance—the liquid rises in the containing vessel when both are similarly heated.

We recognize, therefore, two coefficients of expansion of a liquid. There is first the *coefficient of real expansion*, which is the actual amount by which 1 c.c. of the liquid expands when heated from 0° C. to 1° C. ; and, secondly, there is the *coefficient of apparent expansion*, which is the amount by which 1 c.c. of the liquid expands relatively to the containing vessel when heated from 0° C. to 1° C. Obviously the second coefficient is smaller than the first, and would be zero if the vessel expanded by the same amount as the liquid.

Density and Temperature

In dealing with liquids it is useful to consider the effect of expansion on density, which is, by definition, the mass of unit volume of the liquid. If M , V , and ρ be respectively the mass, volume, and density of a quantity of liquid, then we have the relation expressible in either of the three forms—

$$\left. \begin{array}{l} \rho = \frac{M}{V} \\ V = \frac{M}{\rho} \\ M = \rho V \end{array} \right\} \quad \dots \dots \dots \quad (7.8)$$

Now the mass of a body does not change when it is heated by any amount large enough for us to detect, so that, since the volume changes, the density must change also. Substituting $\frac{M}{\rho}$ for V in equation (7.2), we obtain at once

$$\rho_t = \rho_0 (1 + \beta t) \quad \dots \dots \dots \quad (7.9)$$

which shows how the density of a liquid varies with temperature : obviously it diminishes as the temperature rises.

Determination of β

Let us now consider an ideal experiment, in which a vessel of known volume is first filled with a liquid and weighed

at 0° C. , and then heated to $t^\circ\text{ C.}$ Both vessel and liquid expand, but the greater expansion of the liquid causes some to overflow, as in Fig. 30, leaving the vessel filled at $t^\circ\text{ C.}$ Consider first the apparent expansion; this is simply the

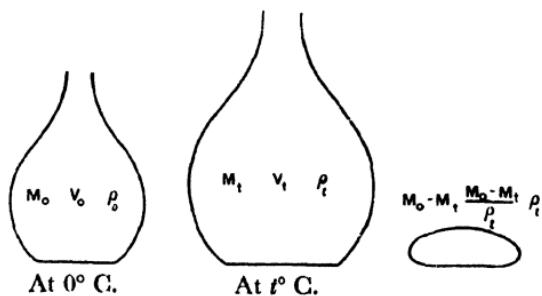


FIG. 30

Thermal expansion of vessel filled with liquid (expansion greatly exaggerated)

amount which has overflowed. If M_0 , V_0 , and ρ_0 be respectively the mass, volume, and density of the liquid filling the vessel at 0° C. , and M_t , V_t , and ρ_t the corresponding quantities at $t^\circ\text{ C.}$ (note that M_t is not the mass at $t^\circ\text{ C.}$ of the liquid which had mass M_0 at 0° C. , but the mass of only a part thereof), the overflow has volume $\frac{M_0 - M_t}{\rho_t}$ and density ρ_t . But if β_1 be the coefficient of apparent expansion, this volume is equal to $V_0\beta_1 t$; i.e. $\frac{M_0\beta_1 t}{\rho_0}$. Hence

$$\frac{M_0 - M_t}{\rho_t} = \frac{M_0\beta_1 t}{\rho_0} \quad \dots \quad (7 \cdot 10)$$

whence, by slight reduction, we obtain

$$\beta_1 = \frac{1}{t} \frac{\rho_0}{\rho_t} \left(1 - \frac{M_t}{M_0} \right) \quad \dots \quad (7 \cdot 11)$$

The real expansion is not only the overflowing portion, but also the amount of liquid filling the additional volume which the vessel has acquired, viz. $V_t - V_0$. Hence the total increase of volume is $\frac{M_0 - M_t}{\rho_t} + (V_t - V_0)$; and since

this must be equal to $V_0\beta t$, where β is the coefficient of real expansion, we have

$$\frac{M_0 - M_t}{\rho_t} + (V_t - V_0) = V_0\beta t \quad \dots \quad (7.12)$$

which gives

$$\beta = \frac{1}{t} \left(\frac{\rho_0}{\rho_t} - 1 \right) \quad \dots \quad (7.13)$$

We can therefore determine the two coefficients by an experiment of this kind if we know the ratio of the densities of the liquid at the two temperatures, for the only other measurements needed are those of M_t , M_0 , and t , which are easily found. It will be noticed that (7.13) is simply another form of (7.9).

Hydrostatic Pressure

The ratio of the densities is best found by a hydrostatic method. To understand this it is necessary to know something about the principles of hydrostatic pressure. At any particular depth below the surface of a liquid at rest there is a certain pressure due to the weight of the liquid above. The pressure is defined, as we know (p. 50), as the force per unit area, and in this case it is the weight of liquid pressing on each unit of area at the depth in question — h , say (together, of course, with the pressure of the atmosphere above). If the liquid is contained in a vessel of cross-sectional area, σ , and if ρ is the density of the liquid, then the mass of liquid above the level considered is $\sigma h \rho$, so that its weight is $\sigma h \rho g$ (equation (1.4)), and

the weight per unit area is therefore $\frac{\sigma h \rho g}{\sigma} = h \rho g$.

Notice that σ has cancelled out, so that the pressure is independent of the shape or size of the vessel used ; for a given liquid at a given place it depends only on the depth below the surface.

Considered as a weight of liquid, the direction of this pressure is, of course, downwards. But by Newton's third law ("To every action there is an equal and opposite reaction"), there is an equal pressure upwards, called into play by the resistance to compression of the liquid beneath. We can see a simple example of this when we place a weight on some soft material, such as wax, resting on a table. The wax is flattened out, but it is flattened at the bottom as much as on the top, although the weight causing the flattening acts downwards only. It calls into play a reaction in the table, and the effect on the wax is as though the table were pressing upwards with the same force as that exerted by the weight downwards.

The piece of wax here considered can be flattened because it is free to spread sideways. A liquid in a vessel with rigid walls, however, cannot so spread. A thin layer at a depth h , subjected to the weight of the liquid above and the reaction of the liquid below, therefore experiences a sideways pressure also, for its tendency to move laterally is resisted by the walls. Considerations such as these show that the same pressure must act in *all* directions at a given depth in a liquid, so that if, for example, a small bladder containing air were placed there, it would be compressed equally all round. This is a general characteristic of hydrostatic pressure, that its value for a given liquid depends only on the depth below the surface, and at each point is the same in all directions.

Measurement of Density Ratios

Now suppose we have a U-tube containing a liquid (Fig. 31). The position of equilibrium is that in which the pressure at

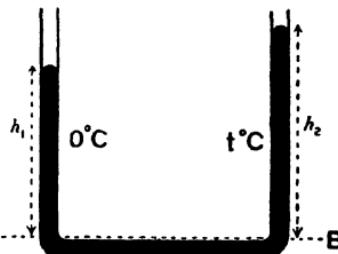


FIG. 31

U-tube containing liquid with limbs at different temperatures

the lowest horizontal level is the same in both limbs, for, if this condition is not satisfied, the greater pressure on one side will force liquid over into the other limb until the pressures are equal. At the level AB, for instance, the pressures in the two limbs are due to the weights of the columns h_1 and h_2 , respectively, of the liquid ; and if ρ_1 and ρ_2 are the densities of the liquid in the two columns, those pressures, as we have just seen, are respectively, $h_1\rho_1g$ and $h_2\rho_2g$. Hence, for equilibrium, $h_1\rho_1g = h_2\rho_2g$. If the same liquid is in both columns, and its temperature is uniform throughout, then $\rho_1 = \rho_2$, so that $h_1 = h_2$; i.e. the liquid stands at the same level in both limbs. If, however, the temperatures on the two sides are respectively 0° C. and t° C., the equation becomes $h_1\rho_0 = h_2\rho_t$, whence

$$\frac{\rho_t}{\rho_0} = \frac{h_1}{h_2} = \frac{h}{h_t} \quad \dots \quad \dots \quad \dots \quad (7.14)$$

if we denote the heights, like the densities, by the corresponding temperature suffixes. We have therefore only to surround the limbs with jackets kept at these temperatures, and read the levels of liquid, in order to determine the ratio of densities.

Experiments on these lines were carried out, with successive improvements in detail, by Dulong and Petit, Regnault, and Callendar and Moss, using mercury, and the coefficients of expansion of this liquid were determined accordingly.

Laboratory Determination of Coefficients

An accurate value having been obtained for the coefficient of real expansion of one liquid, that of another can be found much more simply. Weigh a specific gravity bottle full of mercury, whose coefficient is known, at 0° C. and again at t° C., and let M_0 and M_t be the masses obtained. Let V_0 , ρ_0 , and V_t , ρ_t be the volume and density of mercury filling the bottle at the two temperatures. Now repeat the experiment with the same bottle, but with the liquid under test instead of mercury in it. V_0 and V_t will, of course, have the

same values as before, belonging merely to the bottle, but the masses and densities will be different. Let them be M_0' , ρ_0' , and M_t' , ρ_t' , respectively. We then have

$$\text{and } V_0 = \frac{M_0}{\rho_0} = \frac{M_0'}{\rho_0'} \quad \left. \begin{array}{l} \\ \end{array} \right\} \quad \dots \quad \dots \quad \dots \quad \dots \quad (7.15)$$

$$V_t = \frac{M_t}{\rho_t} = \frac{M_t'}{\rho_t'} \quad \left. \begin{array}{l} \\ \end{array} \right\} \quad \dots \quad \dots \quad \dots \quad \dots$$

$$\text{hence } \rho_0' = \rho_0 \frac{M_0'}{M_0} \quad \left. \begin{array}{l} \\ \end{array} \right\} \quad \dots \quad \dots \quad \dots \quad \dots \quad (7.16)$$

$$\text{and } \rho_t' = \rho_t \frac{M_t'}{M_t} \quad \left. \begin{array}{l} \\ \end{array} \right\} \quad \dots \quad \dots \quad \dots \quad \dots$$

$$\text{but } \rho_0' = \rho_t' (1 + \beta' t) \quad \dots \quad \dots \quad \dots \quad \dots \quad (7.17)$$

where β' is the coefficient of real expansion of the liquid under examination. Hence

$$\rho_0 \frac{M_0'}{M_0} = \rho_t \frac{M_t'}{M_t} (1 + \beta' t) \quad \dots \quad \dots \quad \dots \quad \dots \quad (7.18)$$

$$\text{whence } \beta' = \left(\frac{\rho_0}{\rho_t} \frac{M_0'}{M_0} \frac{M_t}{M_t'} - 1 \right) \frac{1}{t} \quad \dots \quad \dots \quad \dots \quad \dots \quad (7.19)$$

All the quantities on the right-hand side are easily determinable ($\frac{\rho_0}{\rho_t}$, of course, is known from the coefficient for mercury) so that β' can be calculated.

The following are a few values of coefficients of real expansion of liquids. As with solids, they are not quite independent of temperature, and a similar correction term has to be included when great accuracy is required.

Mercury	0.00018
Ethyl Alcohol.	0.00110
Glycerine	0.00053
Ether	0.00163
Water (15° C.)	0.00015

Expansion of Water

The expansion of water is very exceptional, and demands separate consideration. As the temperature falls the coefficient decreases—which is, in fact, not an uncommon occurrence—but as a temperature of about 4° C. is approached it falls off rapidly, and at this temperature vanishes altogether, becoming negative for still lower temperatures. That is to say, when water is cooled below 4° C. it ceases to contract, and begins to expand. Water has thus a maximum density at about 4° C. , as is shown by Fig. 32.

This peculiarity has important effects. When the tem-

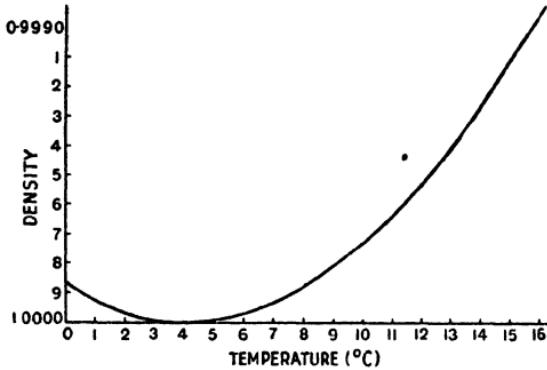


FIG. 32

Density of water between 0°C. and 16°C.

perature of the air falls, the surfaces of ponds and lakes are cooled, and the denser liquid sinks—a process of convection thus being started which tends to equalize the temperature throughout. When the temperature falls to 4° C. , however, further cooling makes the surface water expand and so become less dense. It therefore remains on the surface, and the circulation ceases. It thus occurs that water freezes first on the surface, and since it expands on freezing, the ice so formed floats. Had water followed the usual course of contracting continuously with fall of temperature, ice would have formed at the bottom instead of the top.

EXPANSION OF GASES

Volume and Pressure

In dealing with the thermal expansion of solids and liquids, it was unnecessary to consider other influences than that of rise of temperature, because solids and liquids have such high bulk moduli of elasticity that they are almost incompressible. Gases, however, are very easily compressed by quite small forces, and therefore in measuring their thermal expansion or contraction we must be very careful that their volume is not changed other than by the gain or loss of heat. The general compressibility of gases is so important that it is desirable not to consider it apart from the purely thermal expansion, but to treat the volume, temperature, and pressure of a gas all together, and study their mutual relations.

Pressure of a Gas

Let us first understand clearly what is meant by the pressure of a gas. Imagine a quantity of gas enclosed in a cylinder by a light movable piston which can rise and fall without allowing any leakage at the edges but without any frictional resistance, so that it moves in response to the slightest difference of pressure on the two sides (Fig. 33). If we neglect the weight of the piston, then we have the weight of the atmosphere pressing on it from above, and this, reckoned per unit of area, is the whole pressure on the gas inside the cylinder. Let us suppose the piston remains at rest in the position shown in the figure. Then the pressures on the two sides must be equal, so the gas must exert an upward pressure equal to the weight, per unit area, of the atmosphere.

In terms of the molecular theory, we imagine this to be due to the constant impacts on the piston of the moving molecules of the gas.



FIG. 33

Gas enclosed in cylinder by movable piston

Suppose, now, we lift the piston a short distance. The volume of the gas is then increased, so that the number of molecules per cubic centimetre becomes less, and their impacts are distributed over a greater area of the walls. Fewer impacts are therefore made per second on the piston (*i.e.* the pressure exerted by the gas decreases), and the gas is therefore no longer able to support the pressure of the atmosphere. The piston consequently falls again until the original pressure is reached. If we press the piston down, the reverse happens ; the pressure exerted by the gas increases and pushes it up again. There is only one position in which the piston can remain in equilibrium—namely, that in which the pressures on the two sides are equal.

We shall deal only with gases in equilibrium, and we can therefore speak either of the pressure *on* a gas or the pressure *of* a gas, for they are equal. It must be remembered, of course, that the pressure is the force per unit area, so that in this case it is not the weight of the whole column of air above the piston, but that weight divided by the area of the piston. The same thing, of course, is true of the pressure exerted by the gas.

In dealing with the pressure in a liquid we saw that it depended only on the depth below the surface of the liquid, whereas now that we are speaking of gases we take the pressure to be the same throughout the gas. Strictly speaking, there is no difference at all between the two cases, but the very much smaller density of gases as compared with liquids makes it possible to neglect the variation with depth when we are considering their pressures. To be exact, we should say that the pressure at a depth h below the surface of either a liquid or a gas is the pressure of the atmosphere plus $h\rho g$, where ρ is the density of the gas or liquid. With a liquid the latter term may easily predominate over the former (at a depth of 30 inches, for instance, the pressure due to a column of mercury is equal to that of the whole atmosphere), and it is a common practice to leave the atmos-

pheric pressure out of account when speaking of the pressure in a liquid, although, of course, in certain calculations its value must be included. In others (such as that on p. 114, where the atmosphere presses equally on both limbs of the U-tube) its effect cancels out, and it may be ignored.

With a gas, however, the external pressure, whether that of the atmosphere alone or the atmosphere augmented or diminished by other pressures, is usually so much greater than the pressure due to the weight of the gas itself at the depth considered, that it is the latter term that can be ignored. We therefore speak of the pressure within a gas as being the same throughout the whole volume, and are justified in doing so in all cases that we shall meet with in practice. If, however, we ascend considerable heights into the atmosphere, we experience a diminution of pressure with decrease in depth below the top of the atmosphere, and in such cases the variation can no longer be ignored.

Atmospheric Pressure

The pressure of the atmosphere is measured by a *barometer*, which is a particular form of *manometer*, as any instrument used for measuring pressures is called. In the barometer the atmospheric pressure is, in principle, allowed to push mercury upwards into an evacuated tube, and the pressure is measured by the height to which the mercury can be raised. A clean glass tube closed at one end, of which the cross-section is immaterial but which should not be much less than 1 metre in length, is filled with mercury from which dissolved air has been expelled by boiling, and then inverted over a dish of mercury so that its open end is below the surface of the liquid (see, for example, tube (a) in Fig. 40). The mercury in the tube sinks until its weight per unit area is equal to the atmospheric pressure. For clearly the pressure at the level of the surface of mercury in the dish must be the same inside and outside the tube, otherwise mercury will flow until equality is established. But the pressure

inside is that of the mercury column, and the pressure outside, that of the atmosphere. Hence the height of the mercury measures the atmospheric pressure.

Normal Temperature and Pressure: If h be this height, and ρ the density of mercury, the pressure is gph (p. 112). h varies slightly from day to day through meteorological causes. Its average value is not far from 76 cm., and this is chosen to define the so-called "normal atmosphere." ρ , of course, depends on the temperature, and the normal atmosphere is more strictly defined as the pressure of a column of mercury 76 cm. high at a temperature of 0° C. Air at 0° C. which exerts this pressure is said to be at "normal temperature and pressure"—a phrase usually abbreviated to "N.T.P."

To make the definition of the normal atmosphere quite precise we must state the value of g , which, as we know, varies with terrestrial latitude. The value at latitude 45° is chosen, viz. 980.60 cm. per sec. per sec. Since the density of mercury at 0° C. is 13.596 gm. per cc., the normal atmosphere, measured in dynes per sq. cm., is therefore $980.60 \times 13.596 \times 76 = 1,013,250$. For rough work the pressure of the atmosphere is often taken as a million dynes per sq. cm.

Boyle's Law

These things being understood, let us now see how the pressure, volume, and temperature of a gas are related to one another. Take first of all the special cases in which one of these quantities remains constant while the others vary, beginning with the case of constant temperature. We can decrease the pressure on the gas by pumping away some of the atmosphere over it, or increase the pressure by placing weights on the piston. The result is that in the first case the piston rises to a new equilibrium position (*i.e.* the volume of the gas is increased), and in the second case it falls to a new equilibrium position (*i.e.* the volume of the gas is decreased).

To study this effect quantitatively we use the apparatus

illustrated in Fig. 34. The gas is contained in a bulb attached to a U-tube containing mercury which can be drawn off or to which more mercury can be added at the top of the right-hand limb. If a scale is placed behind each limb, or the limbs are graduated, we can measure the difference in the heights of the mercury, and we will suppose that we know the volume of the bulb and of unit length of the tube. The volume of the gas can then be read off on the left-hand scale, while its pressure is equal to the pressure of the atmosphere plus the amount by which the mercury in the right limb is higher than that in the left. By adding or subtracting mercury we can then vary the pressure, and the volume of gas will change accordingly.

By experiments along these lines it is found that however we change the pressure (provided it does not get extremely great), the product of the pressure P and the volume V remains constant. This is true, whatever gas we take, provided that the temperature remains constant and that the gas is not very near to its condensation point (*i.e.* to the temperature at which it becomes a liquid). There are, it is true, very slight deviations from exactitude in this law, particularly at very high pressures, but these may be ignored for most purposes. The law is known as *Boyle's Law*, and may be expressed by the equation

$$PV = \text{constant} \quad (T \text{ constant}) \quad \dots \quad (7.20)$$

Coefficient of Expansion

Now let us consider how the volume and temperature are related when the pressure is kept constant. This is the ordinary case of expansion already considered for solids and liquids, and we have a similar relation,

$$V_t = V_0 (1 + \beta t) \quad (P \text{ constant}) \quad \dots \quad (7.21)$$

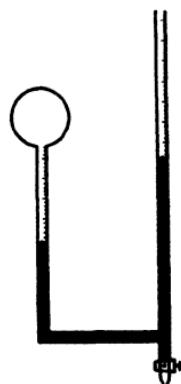


FIG. 34

Apparatus for testing Boyle's Law

where β is the coefficient of expansion of the gas at constant pressure. β can be measured by an apparatus similar to Fig. 34, with the left limb of the tube bent round so that the bulb can be placed in a bath and raised to a known temperature, mercury being added to or subtracted from the right limb in order to keep the pressure constant. Corrections have to be applied for the portion of the gas in the tube which remains outside the bath, and also for the expansion of the bulb itself.

Experiments with various gases show that β has approximately the same value for all gases not too near their condensation points. This value is $\frac{1}{273}$, or 0.00366. Clearly, gases are much more expandible than solids or liquids.

Coefficient of Pressure Increase

If, now, we keep the volume constant and vary the temperature, we find that the pressure changes; the higher the temperature the higher is the pressure. The change follows the same law as the expansion at constant pressure, namely,

$$P_t = P_0(1 + \gamma t) \quad (V \text{ constant}) \quad . \quad . \quad . \quad (7.22)$$

where P_t and P_0 are the pressures at $t^\circ \text{ C.}$ and 0° C. respectively, and γ is approximately a constant, known as the *coefficient of pressure increase*.

We may determine γ by the same apparatus as that used for β , mercury being added or subtracted so as to keep the level constant in the left-hand limb. We then find that, not only is γ approximately constant for all gases (within the same limits as those for β), but that it has the same value as β , viz. $\frac{1}{273}$.

Charles's Law

The statement that every gas expands by the same fraction of its volume when heated at constant pressure from 0° C. to 1° C. is known as *Charles's or Gay Lussac's Law*. Like

Boyle's law, it is approximately true for the so-called "permanent" gases, but not exactly so. An imaginary gas which obeys Boyle's and Charles's laws exactly is called a *perfect* gas, or an *ideal* gas. It is very convenient to assume that a gas is perfect and to calculate its behaviour accordingly, because the laws are then so simple. It must be remembered, however, that the results obtained are only approximately true in practice.

Molecular Theory of Expansion

We can prove that the equality of β and γ is a necessary consequence of Boyle's law, but before doing so let us try to obtain a picture of what is happening in the gas when the temperature changes. Reverting to our cylinder and piston (Fig. 33), suppose first that the pressure is kept constant (e.g. we simply let the atmosphere press on the piston), and that heat is supplied to the gas from below. The temperature rises, so that the molecules move faster. Their impacts on the piston are therefore more vigorous, and they are able to push it up until the decreasing frequency of the impacts, resulting from the greater volume in which the molecules move, makes up for their greater energy. When this happens, the pressure they exert on the under side of the piston is again the same as before, though it is made up of fewer impacts, each with greater energy; and therefore the piston will have reached a new position of equilibrium, higher than the former one, and the volume of the gas is increased.

If, on the other hand, we wish to keep the volume constant as the temperature rises, we must increase the pressure on the piston—by loading it with weights, say—so that the piston remains in its original position. The greater energy of the molecular impacts is then not counteracted by a decrease in the frequency of the impacts, and so the pressure of the gas increases. It increases, of course, by an amount just equal to the additional pressure which we have had to apply to keep the piston from rising. These considerations

give a simple explanation of the general relations between pressure, volume, and temperature of a gas.

Equality of β and γ

To prove that the equality of β and γ follows from Boyle's law, consider a quantity of gas in equilibrium in our cylinder

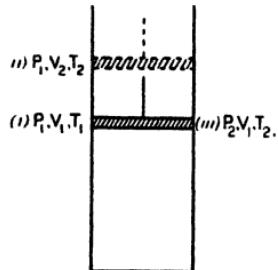


FIG. 35

Variation of pressure, volume, and temperature of a gas

in a state represented by pressure, volume, and temperature equal to P_1 , V_1 , and T_1 respectively. For simplicity let us suppose T_1 is 0° C., and let us call this original state of the gas state (i). Now heat the gas at constant pressure to a state P_1 , V_2 , T_2 , and let us call this state (ii) (Fig. 35). Finally, keeping the temperature constant at T_2 , compress the gas back to its original volume, V_1 . It is then in state (iii)—represented by P_2 , V_1 , T_2 .

Now, comparing states (ii) and (iii), which have the same value of the temperature, we have, by Boyle's law,

$$P_1 V_2 = P_2 V_1$$

$$\text{i.e. } \frac{P_2 - P_1}{P_1} = \frac{V_2 - V_1}{V_1} \dots \dots \dots \quad (7.23)$$

In states (i) and (ii) the pressure is the same, so that

$$V_2 = V_1[1 + \beta(T_2 - T_1)]$$

$$\text{i.e. } \frac{V_2 - V_1}{V_1} = \beta(T_2 - T_1) \dots \dots \dots \quad (7.24)$$

Similarly, comparing states (i) and (iii), in which the volume is the same, we have

$$P_2 = P_1[1 + \gamma(T_2 - T_1)]$$

$$\text{i.e. } \frac{P_2 - P_1}{P_1} = \gamma(T_2 - T_1) \dots \dots \dots \quad (7.25)$$

Hence, substituting from (7.24) and (7.25) into (7.23) we have

$$\beta = \gamma \dots \dots \dots \dots \dots \quad (7.26)$$

The Perfect Gas Equation

The three relations, (7.20), (7.21), and (7.22), in each of which one property of the gas is kept constant, are special cases of one general relation, which may be obtained as follows. Consider a quantity of a perfect gas at 0° C., and at pressure P_0 and volume V_0 . Let it be heated at constant volume to T° C., when the pressure becomes P_T and the volume $V_T = V_0$. We then have, from (7.22),

$$P_T = P_0 (1 + \gamma T)$$

$$\text{i.e. } P_T V_T = P_0 V_0 (1 + \gamma T) \dots \dots \quad (7.27)$$

since $V_T = V_0$.

Now let us choose a new scale of temperature, in which the degrees have the same value as the centigrade degrees, but the zero point is at -273° C. Then temperatures on this scale are obtained simply by adding 273 to the centigrade degrees. Let us denote them by θ , so that

$$\theta = T + 273 \dots \dots \dots \quad (7.28)$$

We then have

$$1 + \gamma T = 1 + \gamma (\theta - 273) = \frac{\theta}{273} \dots \quad (7.29)$$

since $\gamma = \frac{1}{273}$. Hence $P_\theta V_\theta$ (which is simply another way of writing $P_T V_T$) is given by (7.27) as

$$P_\theta V_\theta = P_{273} V_{273} \frac{\theta}{273}$$

$$\text{i.e. } \frac{P_\theta V_\theta}{\theta} = \frac{P_{273} V_{273}}{273} \dots \dots \dots \quad (7.30)$$

Now the right-hand side of this equation is constant, for the numerator is the product of the pressure and volume at a particular temperature, and this is constant by Boyle's

law ; while the denominator is simply a number. Hence the left-hand side is constant also, whatever value θ (and therefore T) might have.

We have supposed here that the same quantity of gas is considered throughout, *i.e.* that its mass is constant. Obviously we can increase the volume of a gas as much as we like without changing the temperature or pressure, if we can add to or subtract from it. If we consider one gram of gas, equation (7.30) may be written

$$PV = R\theta \quad \dots \quad \dots \quad \dots \quad \dots \quad (7.31)$$

where R is known as the *gas constant*. Its value is 83.15×10^6 (when P and V are in C.G.S. units) divided by the molecular weight of the gas. 83.15×10^6 is therefore the value of R for one gram molecule of *any* perfect gas.

Equation (7.31) may be written in another form, by substituting $\frac{M}{\rho}$ for V , in accordance with (7.8). We then obtain

$$M \frac{P}{\rho} = R\theta \quad \dots \quad \dots \quad \dots \quad \dots \quad (7.32)$$

so that if, as before, M is one gram,

$$\frac{P}{\rho} = R\theta \quad \dots \quad \dots \quad \dots \quad \dots \quad (7.33)$$

The meaning of this equation is that if we take one gram of a perfect gas, we can make any two of the three quantities, P , V (or ρ), and θ , whatever we like, but the third will then be determined by the relation expressed by the equation.

Absolute Temperature

Now consider what happens when the temperature of the gas is reduced to -273° C. We have then $\theta = 0$, so that $PV = 0$, and either the pressure or the volume of the gas must vanish. In terms of molecular motions, the vanishing of the pressure means that the molecules must be brought

absolutely to rest, so that they make no impacts on the walls of the containing vessel. The gas may then be said to have absolutely no heat, and for this reason the temperature $- 273^{\circ}$ C. is called the *absolute zero* of temperature, and the θ scale is called the *absolute scale of temperature*.

If we decide to keep the pressure finite and let the volume diminish, then at the absolute zero the volume becomes nothing, *i.e.* the gas disappears. This is never realized in practice, because no gas is perfect: all gases condense into liquids before the absolute zero is reached. The smallest volume a substance can occupy, under terrestrial conditions, is that in which its molecules are in contact with one another, for the molecules themselves are effectively incompressible. At ordinary temperatures the spaces between the molecules of the permanent gases are very large compared with the sizes of the molecules, and if we double the pressure we halve this volume, and so Boyle's law appears to be obeyed. But suppose the molecules are so close together that the spaces between them are comparable with the dimensions of the molecules themselves. Then half the volume of those spaces is no longer approximately half the total volume occupied by the gas (which is, of course, the volume of the molecules and of the intermolecular spaces added together). Boyle's law therefore fails near the condensation point of the gas. A perfect gas is thus seen to be one whose molecules are geometrical points, occupying no space at all. No such gas, of course, can exist.

Gas Thermometers

A gas may be used as a thermometric substance in two ways—in a constant pressure or a constant volume instrument. A perfect gas, of course, would give the same scale with both types, since $\beta = \gamma$ for such a gas. But no gas is perfect, so we have two scales. In practice, however, we usually correct the results, whichever is used, so as to obtain the readings which a perfect gas would give.

EXERCISES

1. What is meant by the coefficient of expansion of a substance? Discuss the differences in the problems of measuring the coefficient for a solid, a liquid, and a gas.
2. A metal rod is found to be 150 cm. long at 20°C . and 150.15 cm. long at 100°C . Find the coefficients of linear and of cubical expansion of the metal.
3. Find the length of a brass rod which will expand by the same amount as a steel rod 1 metre long when heated through the same range of temperature.
4. A barometer with a glass scale reads 762 mm. at 15°C . Find the reading at 0°C . The apparent coefficient of expansion of mercury in glass may be taken as 0.000155 and the coefficient of linear expansion of glass as 0.000089.
5. Explain why the pressure at a point in a liquid is proportional to the depth of the point, and is the same in all directions. Find the pressure at a depth of 1 mile below the surface of the sea, neglecting the compressibility of sea-water and taking its density to be 1.02 gm. per cc.
6. State Boyle's and Charles's laws, and derive the gas equation for a gas which obeys these laws. Show how the behaviour of a perfect gas leads to the idea of an absolute scale of temperature.
7. A litre of dry air weighs 1.293 gm. at N.T.P. At what temperature will a litre weigh 1 gm. at normal pressure, and at what pressure will a litre weigh 1 gm. at normal temperature?

CHAPTER VIII

THE MEASUREMENT OF HEAT

Effects of Heat Transfer

SINCE we cannot observe heat directly, and have no simple means of determining the total kinetic energy of the molecules of a substance, heat is measured in terms of one of its measurable effects. What is usually measured, however, is not the actual heat content of a body, but the amount of heat which it acquires or gives up, and it is an effect of *transfer* of heat that is used to provide the scale of measurement. Three of these effects are of special importance, and each may be used to give us a scale of heat measurement. Only one of them is officially chosen, but the others are so closely related to the measurement that it is best to consider them all together. These three effects of heat transfer are change of volume, change of state, and performance of work.

The Unit of Heat

Change of volume has already been used as a measure of *temperature*, so that measuring heat by the change of volume it produces is equivalent to measuring it by the change of temperature it produces. This is what is actually done. We find that the change of temperature which a body undergoes on receiving heat is directly proportional to the amount of heat, inversely proportional to the mass of the body, and dependent in a not very simple way on the chemical nature of the body. To define a unit of heat, therefore, we choose a particular substance (water), take unit mass of it, and define unit quantity of heat as the quantity which raises the

temperature of unit mass of water by 1° C. This unit is called the *calorie*.

To be precise we have to state what particular degree of temperature we choose, for the amount of heat required is not the same for all degrees. Thus, it takes a different amount of heat to raise 1 gm. of water from 0° C. to 1° C. from what it does to raise it from 15° C. to 16° C. For this reason we speak of the former as the 0° calorie and the latter as the 15° calorie. The *mean* calorie is often used ; it is $\frac{1}{100}$ of the amount of heat required to raise 1 gm. of water from 0° C. to 100° C. For all but the most exact work, however, we may regard the calorie as fairly uniform over moderate ranges of temperature, and disregard its variations.

Change of State

It is supposed here that the water remains water throughout. If we had 1 gm. of a mixture of ice and water at 0° C., and added 1 calorie of heat, the temperature would not rise to 1° C. It would not change at all, but some ice would be melted. In this case the heat is used in changing the state of the substance from solid to liquid. It is found that the mass of ice melted is proportional to the amount of heat added, so that we could have measured heat by saying that the unit of heat is the amount required to melt 1 gm. of ice at 0° C. This would have given us a unit nearly 80 times as large as the calorie.

Latent Heat

It is only for convenience that we speak of the agency which melts ice as "heat." Strictly speaking, we should reserve the name "heat" for that which raises temperature, and say that when heat is supplied to a solid and melts it without change of temperature, the heat changes into something else—called "melting power," say. This would be justified also by consideration of the molecular theory, for we do not give kinetic energy to the molecules when we change ice into

water ; we give them *potential* energy by loosening the bonds between them. Since, however, heat changes so readily into melting power (and is, in fact, recoverable again as heat, for if we put a solid at, say, -10° C. into the mixture, this solid is raised in temperature and water is frozen to ice again), we compromise and call melting power *latent heat*, and define the unit of latent heat as the number of calories required to melt 1 gm. of solid without changing its temperature.

Heat and Work

The third effect of heat—the performance of work—is on a very similar footing. If we supply heat to the gas in the cylinder shown in Fig. 33, we can arrange matters so that it lifts the piston against the pressure of the atmosphere, thus doing work ; and this may be done without the temperature of the gas rising. In this case we do say that the heat is transformed into something else—into work ; we do not speak of “active heat,” as by analogy with latent heat we might legitimately do.

Conservation of Energy

Despite this difference in nomenclature, the relations between heat and latent heat are on exactly the same footing as those between heat and work, though there is one secondary, but practically very important, difference which will be mentioned presently. All three are forms of energy which can be transformed into one another, and we find that in every case when such a transformation takes place, there is a definite and constant ratio between the amount of the one which vanishes and the amount of the other which appears. About 80 calories, for example, are required to melt 1 gm. of ice, and a different (though quite definite and constant) amount to melt 1 gm. of some other substance. Similarly, one calorie produces more than 40 million ergs of work. This quantity is called the *mechanical equivalent of heat*—an exceedingly important quantity in engineering and in physical theory.

We shall return to it later. The law that a given quantity of heat can be transformed into a definite quantity of work is known as the *First Law of Thermodynamics*. It is an example of the still more general law of *conservation of energy*.

We might mention a few examples of these transformations. As we have said, when a hot body is placed in a mixture of ice and water, heat is transformed into latent heat, and if a colder body is then introduced, this latent heat is transformed back into heat which raises the temperature of that body. In general, when we rub two solid bodies together, we do work which is transformed into the heat of friction, and in the steam engine heat generated in the boiler is ultimately transformed into work done by the engine. When we rub together two pieces of ice, however, there is no change of temperature, but the ice melts. Here work is transformed directly into latent heat, and heat properly so-called does not appear at all, since the temperature remains unchanged. On the other hand, when water freezes and bursts our pipes in the winter, latent heat is transformed directly into work. In all these transformations no energy is lost. For every calorie of heat, or unit of latent heat, or erg of work that disappears, an equivalent amount of one of the other forms of energy appears.

Why do we not speak of the "latent heat equivalent of heat" as an analogue of the "mechanical equivalent of heat"? The reason lies in the difference, referred to above, between these two equivalents, which is this. When heat is transformed into work it does not matter what agency effects the transformation; the amount of work produced per unit of heat transformed is always the same, and it is independent of the character of the work done. We do not have to speak of the work of lifting an iron weight as different in magnitude from the work of dragging a wooden log along the ground. But latent heat differs in magnitude for different changes of state. It takes much more heat to change 1 gm. of water into 1 gm. of steam at 100° C. than to turn 1 gm. of ice into

1 gm. of water at 0° C., and a different amount again to melt 1 gm. of lead at its melting point. Thus there is no unique "latent heat equivalent of heat." We should have to mention the particular change of state concerned before such a phrase acquired a definite meaning.

This was not understood by engineers in the eighteenth century. They thought they could get more work from a given quantity of heat by choosing a change of state for which the latent heat was small, and converting the heat first into latent heat and then into work. For example, when a liquid in a cylinder (Fig. 33) is evaporated, the piston is pushed up, thus doing work, and if only a small amount of heat is required to evaporate the liquid, it would seem that the work was being obtained at a cheap rate of expenditure of heat. It was shown by Carnot, however, that this was not so. If you choose a change of state which requires only a small amount of heat, then the latent heat produced gives you only a small amount of work. No matter how you make the change, the amount of work obtainable from the transformation of a given quantity of heat is always the same. All that you are at liberty to do (and it is a very important liberty) is to arrange your transformation apparatus (your heat engine) so that as much of the available heat as possible is transformed into work and not wasted.

Specific Heat

It was said above that the change of temperature which heat produces in a body depends on the chemical nature of the body. This means that the amount of heat required to raise 1 gm. of a substance through 1° C. will not be the same for all bodies. The ratio of this quantity for a given substance to its value for water is called the *specific heat* of the substance. Since the amount of heat required to raise 1 gm. of water through 1° C. is, by definition, 1 calorie, the specific heat of a substance is numerically equal to the number of calories required to raise 1 gm. of it through 1° C. If we call this

quantity S , then clearly the amount of heat required to raise M gm. of the substance from temperature T_1 ° C. to temperature T_2 ° C. is given by

$$H = MS (T_2 - T_1) \quad . \quad . \quad . \quad (8 \cdot 1)$$

and, of course, the same amount of heat is given out if the body falls through this range of temperature.

Heat Capacity

The product of the mass and specific heat of a body is called its *heat capacity*. It is obviously the number of calories required to raise the body through 1° C. The specific heat is characteristic of the *kind* of substance, *e.g.* copper or glass—but the heat capacity is characteristic of a particular sample of that substance. The term “capacity” suggests the old fluid theory of heat, but it is nevertheless sometimes very convenient in picturing what is taking place. For instance, we can get much more heat by allowing a given mass of water to cool through a certain range of temperature than by allowing the same mass of copper to cool through the same range, and we may therefore think of the water as having a greater capacity for heat.

Measurement of Specific Heat

The determination of specific heats is made most conveniently by what is known as the *method of mixtures*, in which two bodies of known mass at different temperatures are made to come to a common temperature by communication of heat from the hotter to the colder. The temperatures are measured, and if the specific heat of one body is known, that of the other can be calculated. Unless the possibility of chemical action or of solution makes it undesirable, it is convenient to take water as one of the substances, since its specific heat is 1. Being a liquid, however, it has to be contained in a vessel, known as a *calorimeter*, and this also takes some of the

heat, which must be allowed for. The calorimeter is usually made of copper—a substance with a good conductivity for heat, so that it quickly acquires the temperature of the water throughout. It is suspended by cotton threads, or supported by cotton wool, in an outer vessel which protects it from draughts and, as far as possible, from loss of heat by conduction and radiation.

Let us suppose we weigh a calorimeter empty, and then about two-thirds full of water. Let the masses be M_c and $M_c + M_w$, so that the mass of the water is M_w . Let the temperature of both be T_1 . Now let a body of mass M , made of the same material as the calorimeter but at a higher temperature, T_3 , be lowered into the water, and let the common temperature attained be T_2 . We then have, by applying equation (8.1) :

$$\text{Heat given out by hot body} = MS (T_3 - T_2)$$

where S is the specific heat of the material of the body. Also

$$\begin{aligned} \text{Heat received by water and calorimeter} &= \\ M.S (T_2 - T_1) + M_w (T_2 - T_1) & \end{aligned}$$

These two quantities are equal, if we neglect the small amount of heat inevitably lost by conduction and radiation during the attainment of equilibrium. We thus have an equation in which the only unknown quantity is S , and S is therefore calculable.

We can now repeat the experiment with the same calorimeter containing water, but with a hot body made of some other substance, of specific heat S' which we require to know. The above expressions still hold good, with S' substituted for S in the formula for the heat given out. S now being known, the equation then serves to determine S' .

In this way we can determine the specific heat of any solid or liquid substance, provided that no change of state or chemical action occurs when the mixing takes place.

Measurement of Latent Heat

Latent heat can be determined in essentially the same way. As an example, suppose the same calorimeter and the same mass of water as before are used at room temperature, and that the substance inserted is a piece of ice of mass M , which is just at its melting point (0° C.). In this case it is the calorimeter and water that give out heat and fall in temperature, and the ice that receives it. The heat it receives is, first, the latent heat required to melt it, and secondly, the heat required to raise the water so formed to the final common temperature, T_2 . If L is the unit of latent heat (usually called simply the "latent heat," for brevity) of fusion of ice, these two quantities are respectively ML and $M(T_2 - 0) = MT_2$. Hence our equation is

$$M_s S (T_1 - T_2) + M_w (T_1 - T_2) = ML + MT_2 . \quad (8 \cdot 2)$$

whence L can be calculated.

In practice M is determined by weighing the calorimeter and contents after the ice has been inserted and the temperature taken, and subtracting $M_c + M_w$. In this way we avoid errors which would arise through premature melting if the ice were weighed beforehand.

The latent heat of vaporization of water can be determined in the same way, by passing steam at 100° C. into the calorimeter instead of ice at 0° C. The value in this case is about 536 calories per gram of steam condensed.

The left-hand side of (8.2) may be written

$$(M_s S + M_w) (T_1 - T_2)$$

The quantity $M_s S$ is called the *water value*, or *water equivalent* of the calorimeter, since the effect of having to use a calorimeter is the same as though the mass of water were increased by this amount and no calorimeter were used. Clearly the water value of the calorimeter is its heat capacity.

Bunsen's Ice Calorimeter: A very accurate method of finding the latent heat of fusion of ice (or, if this is taken as known, the specific heat of some other body) was devised by Bunsen and embodied in his *ice calorimeter*, illustrated in Fig. 36. A is a test-tube surrounded by a vessel B, containing water resting on mercury, which continues into the graduated capillary tube C, open to the air at its end. First a freezing mixture is put in A, so that the outside of that tube is coated with a layer of ice. The freezing mixture is then removed, and the whole apparatus is surrounded by a mixture of ice and water (except that A remains empty and open to the air) until it is uniformly at 0° C. A known mass, M , of a substance of specific heat, S , is then raised to a temperature $T^{\circ}\text{ C.}$ and placed in A. It falls to 0° C. , and the heat it gives out in so doing melts some of the ice in B. The amount melted is indicated by the reading of the mercury in the tube C, for when ice changes into water it contracts by a definite amount—0.0907 cc. for each gram melted.

Suppose the contraction observed is V cc. Then the mass of ice melted must be $\frac{V}{0.0907}$ gm. This requires $\frac{VL}{0.0907}$ calories, which must be supplied by the mass M . The heat given out by the mass M , however, is MST . Hence

$$MST = \frac{VL}{0.0907} \quad \dots \quad (8.3)$$

from which we can find S if L is known or L if S is known.

One advantage of this instrument is that the calorimeter does not change temperature during the experiment, so that

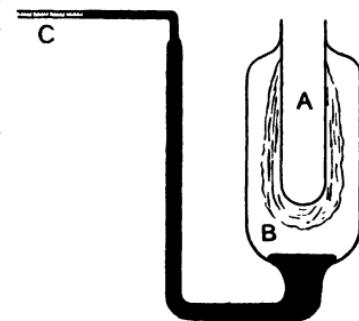


FIG. 36
Bunsen's Ice Calorimeter

its water equivalent is of no importance, and there is no error through loss of heat by conduction or radiation.

Calculation of Mechanical Equivalent of Heat

We have tacitly assumed above that when a substance is heated without change of state, all the heat is used in raising the temperature. This is not exactly true. Bodies expand when heated, and in ordinary circumstances they do so against the pressure of the atmosphere. Hence work is done in pushing back the air, and since bodies have different coefficients of expansion, this work varies from one body to another. With solids and liquids, whose expansion is very small, we can usually ignore this, but with gases the case is different. The specific heat of a gas is very different at constant pressure from what it is at constant volume, for instance, and it is easy to see that the former value is the greater. For, when the volume is constant, no work is done, and all the heat supplied is used in raising the temperature of the gas. At constant pressure, however, some of the heat supplied is used in producing expansion against the external pressure, so that more is required to produce the same rise of temperature.

We can apply these considerations to the determination of the mechanical equivalent of heat. Let us suppose we have measured the specific heats of a gas at constant (normally atmospheric) pressure and constant volume, and let them be S_p and S_v , respectively. Then $S_p - S_v$ is the number of calories required to do the work of expansion. Now if P is the atmospheric pressure and dV is the amount of expansion, this work is PdV ergs. For, suppose the gas is contained in the cylinder shown in Fig. 33. The total force pressing on the piston is $P\sigma$, where σ is the area of the piston ; and the work done in raising the piston through a distance h cm. is therefore $P\sigma h$. But σh is the expansion dV . Hence the work done is PdV .

Now let \mathcal{J} represent the number or ergs of work into

which one calorie of heat can be transformed, *i.e.* the mechanical equivalent of heat. Then $S_p - S_v$ calories become $\mathcal{J}(S_p - S_v)$ ergs, so that

$$\mathcal{J}(S_p - S_v) = PdV. \quad \dots \quad (8.4)$$

Let us take air as the gas under consideration. Experiment gives $S_p = 0.238$ and $S_v = 0.170$. The normal pressure of the atmosphere is 1,013,250 dynes per sq. cm (p. 120), and dV is the expansion of 1 gm. of air when heated through 1°C . One gm. occupies 773.4 cc. at normal pressure, and we know that it expands by $\frac{1}{273}$ of this amount when heated through 1°C . Hence $dV = \frac{773.4}{273}$ cc. Our equation therefore becomes

$$\mathcal{J}(0.238 - 0.170) = \frac{773.4 \times 1,013,250}{273} \quad \dots \quad (8.5)$$

whence $\mathcal{J} = 4.22 \times 10^7$ ergs per calorie. \dots (8.6)

This value is rather larger than that generally accepted as most probable, which is 4.18×10^7 . In the British system of units this corresponds to 777 foot-pounds as the energy required to raise one pound of water through $1^\circ \text{Fahrenheit}$.

The Principal Specific Heats

A gas may, of course, absorb or give out heat without either the pressure or the volume remaining constant. The specific heat will then depend on the circumstances of the process. The specific heats at constant pressure and at constant volume are, however, the most important, and they are known as the *principal* specific heats of a gas. Their accurate determination is a difficult matter. The specific heat at constant pressure is best determined by a method devised by Callendar and illustrated diagrammatically in Fig. 37. A steady stream of the gas is passed along a tube in which is a coil of wire, C, heated electrically by a steady current. The temperature of the gas is observed by thermometers at

entrance (T_1), and at exit (T_2). All the conditions being quite steady, the heat generated in the wire per second is equal to that received by the gas per second, provided that none is lost by radiation or in other ways. This is ensured as far as possible by surrounding the apparatus by a vacuum jacket, and that in turn by a jacket of water at constant

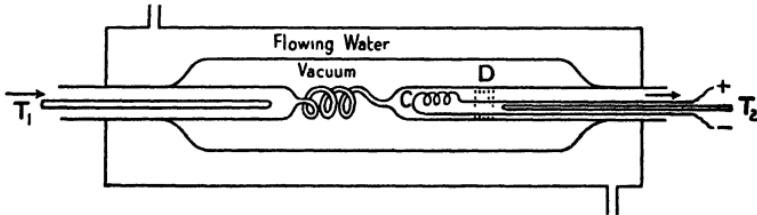


FIG. 37

Steady Flow Method of determining Specific Heat of a gas at constant pressure

temperature. A set of wire gauze screens, D , serves to mix the gas before the final temperature is taken.

The heat generated in the wire is given by $VA \times 10^7$ ergs, where V is the potential difference, in volts, between the ends of the wire, and A amperes is the current. The rate at which the gas flows must be determined, and this gives the mass, M , passing per second, while the difference of the two thermometer readings gives the rise of temperature produced. We then have the equation

$$VA \times 10^7 = \mathcal{J}MS, (T_2 - T_1) \dots \dots \quad (8.7)$$

which gives us S , if \mathcal{J} is known.

The same method has been applied to the measurement of the specific heat of water at various temperatures.

Experimental Determination of \mathcal{J}

The foregoing method, of course, cannot be used unless we can determine \mathcal{J} by some other means ; and the calculation of \mathcal{J} already made presupposes a knowledge of S , so we still need an independent determination of \mathcal{J} . This can be

obtained by experiments of various types. Joule, for example, in a very elaborate and accurate experiment, allowed falling weights to turn paddles in a calorimeter of water, in which case the energy of the weights was converted into heat in the water by friction. Alternatively, we can determine S_p by the method of mixtures, passing the gas into a calorimeter and noting the rise of temperature produced. This was done, with many precautions, by Regnault. Callendar's constant-flow method may be regarded as the best method of determining the product $\mathcal{J}S_p$, so that the more accurately we can determine either of these quantities, the more accurately we can know the other.

Specific Heat at Constant Volume

The specific heat of a gas at constant volume may be determined by *Joly's steam calorimeter*, in which two similar spheres containing different amounts of the gas are hung from opposite arms of a balance which is brought to equilibrium by the addition of weights where necessary. They are both raised to 100° C. by the prolonged circulation of steam around them. While their temperature is rising, steam condenses on the spheres, but more condenses on the sphere containing the greater quantity of gas because there is more substance to be heated so that more latent heat must be released. The weights necessary to restore equilibrium measure the excess of steam condensed. If this be m gm. and the latent heat of condensation be L calories, then mL calories are required to raise the excess of gas from the original temperature to 100° C. Hence the specific heat can be calculated, and the volume has been kept constant, apart from expansion of the spheres, which must be corrected for.

In practice, the specific heat of a gas at constant volume is more often determined from the value for constant pressure together with the value of the ratio, $\gamma = \frac{S_p}{S_v}$, which can be determined independently.

Internal and External Work

The determination of γ from the difference of the specific heats is subject, of course, to whatever errors exist in the evaluation of the specific heats, and also depends on the assumption that all the heat which does not raise the temperature of the gas is used in pushing back the atmosphere, or doing *external work*, as it is called. This is true only if no work is done in merely changing the volume of the gas ; or, in terms of the molecular theory, if the molecules of a gas have no attraction for one another which would require the expenditure of work (called *internal work*) to separate them.

Special experiments have been made to test this, in the most accurate of which (first performed by Joule and Thomson) the gas is forced at a steady rate through a throttle, in the form of a *porous plug*, to a region of lower pressure where it can expand. If work is necessary to separate the molecules, the gas should have a lower temperature after expansion, for the work can be done only at the expense of the molecular kinetic energy. If, on the other hand, the molecules should have a natural *repulsion* for one another, a *rise* of temperature would occur, for they would have opportunity to convert their potential energy into kinetic energy when less pressure is exerted on them.

In experiments on these lines, all the gases tested, except hydrogen, showed a slight fall of temperature, indicating a very slight attraction between the molecules. With such gases, then, some work is done simply by enlargement of volume, even when there is no external pressure. It is, however, very slight and can usually be ignored, although, as we shall see (p. 177) it has been put to important uses. In hydrogen a very slight *rise* of temperature occurred, but under the conditions of the experiments, this might have been due to other causes than repulsion between the molecules. All that we can safely infer is that if any work is needed to separate the molecules of hydrogen, it is extremely small.

In solids and liquids, of course, where the molecules are practically in contact with one another, there is strong attraction, as we know from their cohesion. But the phenomena of surface tension suggest that the attraction falls off very rapidly as the distance between the molecules increases, and in gases at ordinary pressures this distance is very large compared with the size of the molecules.

EXERCISES

1. Discuss the relations between heat, work, and latent heat, and give examples of the transformation of energy from one of these forms to another.
2. What is meant by *specific heat* and *heat capacity*? 40 gm. of a solid at 15° C. are placed in a calorimeter of the same material containing 55 gm. of a liquid at 60° C. , and a common temperature of 45° C. is reached. The mass and water value of the calorimeter are respectively 22 gm. and 2.2 gm. Find the specific heat of the liquid.
3. A piece of ice, initially at -50° C. , is supplied with heat at a constant rate until the temperature reaches 150° C. , the pressure remaining constant throughout at 1 atmosphere. Draw and explain graphs showing qualitatively the changes in volume and temperature during the process, and indicate the changes of state which occur.
4. 15 gm. of molten paraffin wax at 70° C. are poured into a copper calorimeter containing water at 10° C. , the weight of the calorimeter being 70 gm. and the weight of the water 50 gm. The final common temperature is 18.6° C. Find the latent heat of fusion of the wax, if the specific heats of copper, liquid wax and solid wax are respectively, 0.095, 0.61, and 0.43, and the melting point of wax is 55° C.
5. What is meant by *internal work* and *external work*? Find the amount of external work done when 1 litre of water at

0° C. freezes to ice at the same temperature and at normal atmospheric pressure. The density of water at 0° C. is 0.99988, and that of ice is 0.91674.

6. Define the "mechanical equivalent of heat." If 8,000 calories are produced by the consumption of 1 gm. of coal, and an engine uses one-tenth of the energy produced by the combustion of coal in pumping water against gravity, find the amount of coal required to enable it to raise 10,000 litres of water to a height of 50 metres.
7. Describe methods for measuring the two principal specific heats of a gas. Why do these specific heats differ from one another?

CHAPTER IX

RELATIONS BETWEEN HEAT AND WORK IN A GAS

GENERAL PRINCIPLES

LET US consider the transformation of heat into work, and work into heat, as exemplified by the behaviour of a perfect gas. We have no changes of state to consider, so that no latent heat is in question, and we suppose, further, that no internal work is done, *i.e.* that the molecules of a perfect gas not only have no size but also exert no forces on one another. Our gas may receive or lose heat, perform work or suffer work to be done on it, and it may do work at the expense of its own heat, but we suppose that no other change occurs. The results we obtain will be applicable approximately to actual gases under less restricted conditions, so that they are not merely mathematical exercises, but we ignore that for the moment. We suppose throughout that we are dealing with one gram of gas.

We have four principles to guide us :

(i) When heat, H , is transformed into work, W , or work, W , into heat, H , the following relation holds good :

$$W = \mathcal{J}H. \dots \dots \dots \dots \quad (9 \cdot 1)$$

(ii) The state of the gas is determined completely and uniquely by the three quantities P , V , and θ —the pressure, volume, and absolute temperature. Actually, only two of

these are necessary, for if any two of these are given, the value of the third necessarily follows from the relation

$$PV = R\theta. \dots \dots \dots \quad (9.2)$$

(iii) The criterion for a change in the heat content of the gas is that its temperature changes. Hence, if the gas gains or loses heat, the amount of heat gained or lost is given by $\pm MS(\theta_2 - \theta_1)$, where S is the specific heat appropriate to the conditions of the change (e.g. constant pressure, or constant volume), and θ_1 and θ_2 are the initial and final temperatures respectively. Since we are dealing with unit mass of gas, the factor $M = 1$, and since it is convenient to consider only small changes at a time, we can put $\theta_2 - \theta_1 = d\theta$. Our expression for the heat gained or lost is therefore

$$H = \pm Sd\theta. \dots \dots \dots \quad (9.3)$$

(iv) The criterion for the performance of work by or on the gas is that its volume changes. If P is the pressure and $\pm dV$ the change of volume, the work, W , done is

$$W = \pm PdV. \dots \dots \dots \quad (9.4)$$

in which a positive value of dV denotes work done by, and a negative value work done on, the gas. This has already been proved (p. 138) for a gas in a vertical cylinder, but it is true generally, whatever the shape of the containing vessel.

CHANGE OF CONDITION OF A GAS

With the aid of these four principles, let us consider a change from a condition represented by P_1 , V_1 , θ_1 to a condition represented by P_2 , V_2 , θ_2 . Let us suppose the change is brought about by the addition of heat, and calculate the heat necessary to make it. Since heat and work are forms of energy, which cannot be destroyed or created, and since

the values of P , V , and θ determine the condition of the gas uniquely, the amount of heat we are seeking is the same by whatever series of processes the change is made. Let us suppose, then, that it is made in two steps. First, let the gas be heated at constant volume until its temperature is θ_2 ; the pressure will then have some value, P_1 , say. Second, let the volume change to V_2 at constant temperature. The pressure must then be P_2 , and the required change is complete.

The heat taken in during the first step is $\int_{\theta_1}^{\theta_2} S_v d\theta$. In the second step the gas neither acquires nor loses heat, since its temperature remains constant. It does work, however, and since it does not draw on its own heat for the purpose, heat must be supplied from the outside. The work done is $\int_{V_1}^{V_2} P dV$, so that the heat supplied is $\frac{1}{J} \int_{V_1}^{V_2} P dV$. The total amount of heat supplied is therefore

$$H = \int_{\theta_1}^{\theta_2} S_v d\theta + \frac{1}{J} \int_{V_1}^{V_2} P dV. \quad \dots \quad (9.5)$$

If we assume S_v to be the same at all temperatures, the first integral is simply $S_v(\theta_2 - \theta_1)$. To evaluate the second, we put $P = \frac{R\theta_2}{V}$ from (9.2), since the temperature is constant at θ_2 throughout this change. The integral is then $\frac{R\theta_2}{J} \int_{V_1}^{V_2} \frac{dV}{V}$, i.e. $\frac{R\theta_2}{J} \log \frac{V_2}{V_1}$. Hence

$$H = S_v(\theta_2 - \theta_1) + \frac{R\theta_2}{J} \log \frac{V_2}{V_1}. \quad \dots \quad (9.6)$$

Now let us suppose the same change to be made by another process. Let the temperature of the gas be first

changed to θ_2 at constant *pressure*; and then let the pressure be changed to P_2 at constant temperature. The volume must then, of course, be V_2 . The heat taken in at the first step (again assuming the specific heat to be independent of temperature) is $S_v(\theta_2 - \theta_1)$, and that at the second step is $\frac{1}{J} \int_{P_1}^{P_2} P dV$.

To evaluate the second integral we change the variable from V to P . Differentiating equation (9.2) we have, at constant temperature,

$$PdV + VdP = 0 \quad \dots \quad \dots \quad \dots \quad (9.7)$$

so that the integrand is $-VdP = -\frac{R\theta_2}{J} \frac{dP}{P}$. The integral

is therefore $-\frac{R\theta_2}{J} \int_{P_1}^{P_2} \frac{dP}{P} = -\frac{R\theta_2}{J} \log \frac{P_2}{P_1}$. Hence

$$H = S_v(\theta_2 - \theta_1) - \frac{R\theta_2}{J} \log \frac{P_2}{P_1} \quad \dots \quad \dots \quad (9.8)$$

Isometric, Isopiestic, and Isothermal Changes

The expressions (9.6) and (9.8) (which, of course, are alternative expressions for the same amount of heat) may now be applied to a few special cases. If, for instance, $V_2 = V_1$, so that the volume of the gas remains constant, the second term of (9.6) vanishes, and we have

$$H = S_v(\theta_2 - \theta_1) \quad \dots \quad \dots \quad \dots \quad (9.9)$$

This is called an *isometric* change.

If, on the other hand, $P_2 = P_1$, so that the pressure of the gas remains constant, the second term of (9.8) vanishes, and we have

$$H = S_v(\theta_2 - \theta_1) \quad \dots \quad \dots \quad \dots \quad (9.10)$$

This is called an *isopiestic* change.

These results, of course, are immediately obvious. If, however, $\theta_2 = \theta_1 = \theta$, so that the *temperature* of the gas remains constant, then the first terms of both (9.6) and (9.8) vanish, and we have

$$H = \frac{R\theta}{J} \log_e \frac{V_2}{V_1} = - \frac{R\theta}{J} \log_e \frac{P_2}{P_1} \dots \dots \quad (9.11)$$

This last change is called an *isothermal* change.

In an isothermal change the heat taken in, which is given by either of the expressions in (9.11), is used up entirely in doing external work, so that none is left to heat the gas.

Meaning of the Gas Constant

We obtain a very important result by equating the two expressions (9.6) and (9.8) for H . Let us suppose the change is small, so that $\theta_2 - \theta_1 = d\theta$, $V_2 - V_1 = dV$, and $P_2 - P_1 = dP$. Then we have

$$H = S_p d\theta + \frac{R\theta}{J} \log_e \left(1 + \frac{dV}{V} \right) = S_v d\theta - \frac{R\theta}{J} \log_e \left(1 + \frac{dP}{P} \right) \quad (9.12)$$

$$\begin{aligned} \text{i.e. } (S_p - S_v) d\theta &= \frac{R\theta}{J} \log_e \left[\left(1 + \frac{dV}{V} \right) \left(1 + \frac{dP}{P} \right) \right] \\ &= \frac{R\theta}{J} \left(\frac{dV}{V} + \frac{dP}{P} \right) \end{aligned} \quad \dots \quad (9.13)$$

by applying the logarithmic theorem, $\frac{dV}{V}$ and $\frac{dP}{P}$ being very small. Substituting PV for $R\theta$, we have then

$$(S_p - S_v) d\theta = \frac{1}{J} (P dV + V dP) = \frac{R d\theta}{J} \quad \dots \quad (9.14)$$

whence

$$J (S_p - S_v) = R \quad \dots \quad (9.15)$$

This gives us a physical meaning for the gas constant R . It is the value, in work units, of the difference between the principal specific heats of a gas. In other words, it is the work

done when unit mass of a perfect gas is heated through 1° C. at constant pressure.

Adiabatic Changes

We shall apply equations (9.6) and (9.8) to one other case, namely, that in which no heat is supplied to the gas, but yet its condition changes from P_1, V_1, θ_1 to P_2, V_2, θ_2 . This can happen only if an interchange between heat and work takes place *inside* the gas ; as, for instance, when it does work at the expense of its own heat. Such a change is called an *adiabatic* change. It is one in which no heat crosses the boundary of the gas, but not, of course, a change in which the heat content of the gas remains constant.

We can determine the circumstances of such a change by putting $H = 0$ in equation (9.12), when we obtain

$$\left. \begin{aligned} 0 &= S_v d\theta + \frac{R\theta}{J} \frac{dV}{V} \\ &= S_v d\theta + \frac{PdV}{J}, \text{ from (9.2)} \end{aligned} \right\} \quad \dots \quad \dots \quad \dots \quad \dots \quad (9.16)$$

Now, by differentiating (9.2) we see that

$$d\theta = \frac{1}{R} (PdV + VdP) \quad \dots \quad \dots \quad \dots \quad \dots \quad (9.17)$$

and substituting this, and the value of R given by (9.15), in (9.16), we obtain

$$0 = \frac{S_p}{J(S_p - S_v)} (PdV + VdP) + \frac{PdV}{J} \quad \dots \quad (9.18)$$

which reduces to

$$\frac{S_p}{S_v} PdV + VdP = 0. \quad \dots \quad \dots \quad \dots \quad \dots \quad (9.19)$$

The ratio of the principal specific heats of a gas, $\frac{S_p}{S_v}$, is always denoted by γ . Using this symbol, and integrating (9.19), we obtain

$$PV^\gamma = \text{constant.} \quad \dots \quad \dots \quad \dots \quad \dots \quad (9.20)$$

This is the very important equation to an adiabatic change. In such a process, P , V , and θ all change, but in such a manner that this relation is obeyed throughout. We can express it in terms of P and θ or V and θ if we wish, by eliminating the undesired variable between this equation and (9.2). The results—which are, of course, equivalent to (9.20)—are

$$\theta V^{\gamma-1} = \text{constant} \quad \dots \dots \dots \quad (9.21)$$

$$\text{and } P^{1-\gamma} \theta^\gamma = \text{constant} \quad \dots \dots \dots \quad (9.22)$$

Ratio of the Principal Specific Heats

The relation just obtained enables us to find the ratio of the principal specific heats directly. There are various methods of doing this, one of the most interesting being that based on measurements of the velocity of sound. Sound waves consist of alternate compressions and rarefactions of the medium (usually air) in which the sound travels, and these take place so rapidly that there is no time for heat to be taken in or given out during their occurrence, so that they constitute adiabatic changes. In these circumstances it can be shown (see Chapter XIV.) that the velocity of the waves is $\sqrt{\frac{\gamma P}{\rho}}$, where P is the pressure and ρ the density of the air.

Taking the experimental value of the velocity, viz. about 33,200 cm. a second, we find $\gamma = 1.4$, which agrees with the value found from direct measurements of the specific heats.

The value of γ varies for different gases ; it depends on the complexity of the molecules. Thus, for monatomic gases, like mercury vapour and the noble gases, its value is 1.66, while carbon dioxide, a triatomic gas, has a value of about 1.3. Oxygen, nitrogen, and hydrogen show the same value as air.

We see by equation (9.21) that an adiabatic change involves a change of temperature. A particular example, of frequent occurrence, is the rise produced when a gas is suddenly compressed, or the fall of temperature when it

suddenly expands. The change in such a case is not always adiabatic, since it is difficult to prevent heat from being conducted into or away from the gas, but it may be so. This change of temperature may be very great. The heat produced when a bicycle tyre is rapidly pumped up is a common example.

The Indicator Diagram

A diagram in which the pressure of a gas is plotted against its volume is called an *indicator diagram*. In Fig. 38, AB represents an isothermal change on such a diagram. Since the equation

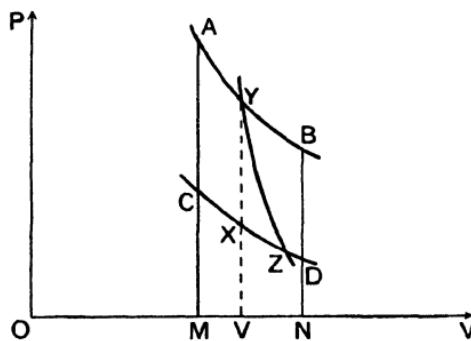


FIG. 38

The Indicator Diagram
AB, CD Isothermals
YZ Adiabatic

to this curve is $PV = \text{constant}$, it is clearly a rectangular hyperbola. CD is another isothermal at a lower temperature than AB, for if we take equal volumes OV, the pressure, VX , is less on CD than that, VY , on AB, so that the temperature must be less. If we suppose that the gas, in the condition represented by Y, is suddenly allowed

to expand adiabatically, its temperature will fall, as equation (9.21) shows. When it has fallen to the temperature corresponding to CD, the change is represented by such a course as YZ, which clearly must be steeper than the isothermal AB, since it has to meet the lower isothermal.

The work done when the gas changes isothermally from a condition $P_1 V_1$, represented by A, to a condition $P_2 V_2$, represented by B (given, by (9.11), as $R\theta \log \frac{P_2}{P_1}$) is $\int_{V_1}^{V_2} P dV$.

This is the area MABN. Hence the work done in an isothermal change is represented on the indicator diagram by the area under the curve representing the change. The same result holds, of course, for any other change, e.g. an adiabatic change—for the work in every case is $\int_{V_1}^{V_2} P dV$.

EXERCISES

1. Define *isothermal* and *adiabatic* changes in a gas. If the pressure of a perfect gas is doubled, find the relative change of volume when the change is made (a) isothermally ; (b) adiabatically.
2. Find the work done when 1 gram molecule of a perfect gas expands to half the original pressure isothermally at 0° C. ($R = 83.15 \times 10^6$ ergs per gm. mol.)
3. One litre of a perfect gas at 0° C. is compressed adiabatically to 700 cc. Find the change of temperature. If the pressure was originally 1 atmosphere, what is the final pressure?

CHAPTER X

THE SECOND LAW OF THERMODYNAMICS

Reversible and Irreversible Changes

AN important distinction between the changes that can take place in a gas (and indeed in any system of bodies) is that between reversible and irreversible changes. We have seen that when a hot body is placed in a mixture of ice and water it falls in temperature and some ice melts. This change is irreversible ; the latent heat will not transform itself back into heat which raises the body introduced to its original temperature again. The latent heat, it is true, may be transformed back into the same quantity of heat if we insert a *colder* body, but that is not a reversal of the original process.

Reversible processes are rare. An example is the gradual change of the pressure of a gas—contained in the cylinder of Fig. 33, for example. If weights are slowly added to the piston the gas is compressed, and when the weights are removed it returns to its original volume. Here the original state of affairs is recovered in every respect. The change must be very slow, however, otherwise the compression will generate heat which will be conducted or radiated away, and the processes of conduction and radiation are irreversible.

It is, in fact, the spontaneous tendency of heat to pass from a body of higher to a body of lower temperature that makes a truly reversible process an unattainable ideal. Heat flows naturally in one direction and not in the opposite direction, so that whenever the uncontrolled passage of heat occurs, it cannot be reversed without the expenditure of some effort ; and if it is so reversed we have not recovered the original

circumstances because the energy expended in the effort is no longer available. It can be shown—and, indeed, it seems almost obvious—that if we want to make a transformation of energy from one form to another, the most efficient process (*i.e.* the process in which there is least waste) is the one which is most nearly reversible, for it is only waste of energy in the process that prevents reversal. The practical problem is always to convert energy from other forms into work, so that if we wish to convert heat into work—as we do in our heat engines—we must look for a process as nearly reversible as possible.

Carnot's Ideal Engine

An ideal engine, which is quite reversible, was imagined more than one hundred years ago by Sadi Carnot. Suppose the cylinder in Fig. 33

can be made perfectly conducting or non-conducting at will, and, keeping the temperature constant, let us supply the substance in it with a quantity of heat, H_1 , in such a manner that the process could be reversed at any moment. The piston will rise, of course, doing the work shown in Fig. 39 by the area ABNM. Now stop the supply of heat and let the piston rise again adiabatically by reduc-

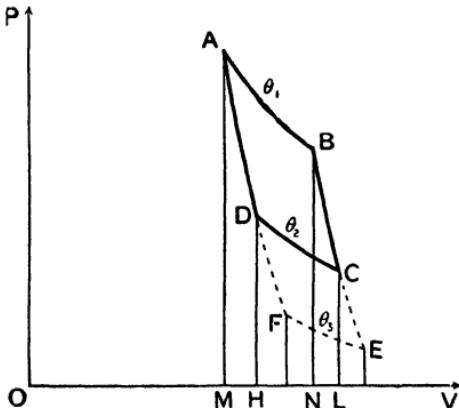


FIG. 39

The Carnot Cycle

AB, CD Isothermals

BC, DA Adiabatics

ing the pressure. This change is represented in Fig. 39 by BC, and work BCLN is done by the gas. Now compress the gas isothermally to the point D, checking the tendency for the

temperature to rise by drawing off heat H_2 . In this process a quantity of work DCLH is done *on* the gas—the work value, of course, of the heat H_2 . Finally, let the gas be compressed adiabatically to A again by doing work DHMA on it.

Now if this process is to be reversible, no heat must have been lost by conduction or radiation. We may in imagination eliminate radiation by supposing that the surroundings are always at the same temperature as the cylinder, and in order to prevent conduction the heat taken in and given out during the isothermal changes must be supposed to pass from and to bodies at the same temperature as the cylinder. Thus, if the temperature along AB be θ_1 , and that along CD be θ_2 , the heat H_1 must be supplied from a source at θ_1 , and the heat H_2 must be given to a sink (the *refrigerator*, as it is usually called) at θ_2 . During the adiabatic changes, of course, no heat passes at all. If these conditions are fulfilled we can take the contents of the cylinder backwards through the series of changes to the starting point, and everything will be in its original state.

In this process along the path ABCDA, the changes that have occurred are (1) a quantity of heat, H_1 , has been taken from a source at a temperature θ_1 , and a quantity H_2 has been given to a refrigerator at a temperature θ_2 ; (2) a net amount of work represented by the area ABCD has been done by the contents of the cylinder; it can easily be seen that ABCD represents this work (which we will call W_1), since it is the difference between the areas indicating work done by and work done on the substance. Now, since the substance has been restored to its original condition, this work can only have come from the heat not given up to the refrigerator, viz. $H_1 - H_2$. Hence

$$W_1 = \mathcal{J}(H_1 - H_2). \quad \dots \quad (10-1)$$

Efficiency

The ratio of the work done to the heat taken in from the source (measured in work units) is called the *efficiency* of the

engine. It is clearly $\frac{W_1}{JH_1}$. It can be shown that the efficiency of the Carnot cycle is the maximum possible for any engine working between the temperatures θ_1 and θ_2 .

No assumption has been made concerning the nature of the substance in the cylinder, which may, indeed, be anything at all. The argument is concerned only with the heat taken in and given out, and the work done. The efficiency of the Carnot cycle, then, is independent of the nature of the "working substance" of the engine. Of course, in practice it would be more convenient to put a gas than, say, a collection of guinea-pigs in the cylinder, but the same principles hold good in the two cases.

The Kelvin Temperature Scale

We have placed no restriction on the amounts of heat taken in and given out, or the particular temperatures at which the changes take place. Let us, then, imagine the process repeated, with the lower isothermal chosen as EF (temperature θ_3) instead of CD (Fig. 39). We shall then have a similar equation,

$$W_1' = J(H_1 - H_3) \dots \dots \quad (10 \cdot 2)$$

where W_1' is represented by the area ABEF, and H_3 is the heat given out along EF. Lord Kelvin proposed a new scale of temperature, such that

$$\frac{\theta_1 - \theta_3}{\theta_1 - \theta_2} = \frac{W_1'}{W_1} \dots \dots \quad (10 \cdot 3)$$

That is to say, the difference of temperature between two bodies is *defined* as being proportional to the amount of work performed when a Carnot engine takes a fixed quantity of heat from the hotter and gives heat to the colder. The advantage of this scale is that it is independent of the properties of any particular substance. It is therefore called the *absolute*, or *thermodynamic* or *Kelvin* scale of temperature.

We have already met with an "absolute" scale, defined in terms of the behaviour of a perfect gas. Although the Kelvin scale is derived in a quite independent way, it can be shown that it is identical with this scale, so there is no confusion in the use of the term "absolute" for both. We may express the position in this way. A difference of temperature presents us with two possibilities (among others): we may produce an expansion (or pressure increase) of a substance when it changes from one temperature to the other; or we may make a substance do work by falling from one temperature to the other. If we choose the latter effect as a measure of the temperature difference it does not matter what substance we choose as a vehicle for the transformation; so long as the transformation takes place by means of a Carnot cycle, the amount of work produced per unit of heat is the same for all substances. If, however, we choose expansion or pressure increase, we get different effects for different substances, and only if the substance is a perfect gas do we get a scale agreeing with the work scale.

It will be noticed that in either case the scale is an ideal one, not directly applicable in practice. The work scale is inapplicable because the Carnot engine cannot be realized, and the perfect gas scale because no gas is perfect. Actual thermometers must be made of actual substances, and their readings may then be corrected to the absolute scale from our knowledge of the behaviour of the substances.

The Absolute Zero

It is clear from Fig. 39 that the lower the temperature of the refrigerator, the greater is the amount of work obtained from the heat H_1 . Thus, when working between θ_1 and θ_2 , the work done is ABCD, and heat H_2 is given to the refrigerator. When working between θ_1 and θ_3 , however, the work done is the larger quantity ABEF, and so the heat H_3 now given to the refrigerator must be less than H_2 . Let us next suppose the adiabatic BC continued down to such a temperature that

all the heat H_1 is converted into work. Then no lower temperature is possible, since, if it were, we could get more work out of the heat H_1 than its equivalent, $\mathcal{J}H_1$, and this would violate the First Law of Thermodynamics. Hence we have reached an absolute zero of temperature. This absolute zero can be shown to be the same as that already defined (p. 127) in terms of the perfect gas scale.

Entropy

Now let us suppose that a Carnot engine works round the cycle ABCDA, taking in heat H_1 at θ_1 , and giving out heat H_2 at θ_2 . Then equation (10.1) holds good, viz.,

$$W_1 = \mathcal{J}(H_1 - H_2) \quad \dots \quad \dots \quad \dots \quad (10.1)$$

and on the Kelvin scale of temperature,

$$\theta_1 - \theta_2 = kW_1 \quad \dots \quad \dots \quad \dots \quad (10.4)$$

where k is a constant. Next, let the engine again take up this quantity of heat, H_2 , at θ_2 , and give out H_3 at the lower temperature, θ_3 , performing work W_2 , so that

$$W_2 = \mathcal{J}(H_2 - H_3) \quad \dots \quad \dots \quad \dots \quad (10.5)$$

We have then

$$\theta_2 - \theta_3 = kW_2 \quad \dots \quad \dots \quad \dots \quad (10.6)$$

Let this process be continued, the heat given out at each stage being taken up again at the next and partly converted into work, until the absolute zero is reached. We shall then have finally

$$W_n = \mathcal{J}H_n \quad \dots \quad \dots \quad \dots \quad \dots \quad (10.7)$$

$$\text{and } \theta_n = kW_n \quad \dots \quad \dots \quad \dots \quad \dots \quad (10.8)$$

since H_{n+1} and θ_{n+1} will be zero. Add together all these quantities of work, W_1, W_2, \dots, W_n . Then, from equations

(10.1), (10.5), (10.7), and the similar equations for the intermediate steps, we see that

$$W_1 + W_2 + \dots + W_n = \mathcal{J}H_1 \dots \quad (10.9)$$

and also, by adding together (10.4), (10.6), (10.8), and the similar equations for the intermediate steps, we obtain

$$\theta_1 = k(W_1 + W_2 + \dots + W_n). \quad \dots \quad (10.10)$$

Hence, from (10.9) and (10.10),

$$\frac{H_1}{\theta_1} = \frac{1}{\mathcal{J}}. \quad \dots \quad \dots \quad \dots \quad \dots \quad (10.11)$$

Now the right-hand side of this equation is constant. Hence the left-hand side is constant also—that is to say,

$$\frac{H_1}{\theta_1} = \frac{H_2}{\theta_2} = \dots = \frac{H_n}{\theta_n}. \quad \dots \quad \dots \quad \dots \quad (10.12)$$

The quantity $\frac{H}{\theta}$ is very important ; it is called *entropy*. A body absorbing a quantity of heat H at a temperature θ absorbs entropy $\frac{H}{\theta}$, and a body giving out heat H at a temperature θ loses entropy $\frac{H}{\theta}$. It is a characteristic of the Carnot cycle that it involves no change of entropy. For we have just proved that, for the isothermals, the entropy $\frac{H_1}{\theta_1}$ gained along AB is equal to the entropy $\frac{H_2}{\theta_2}$ lost along CD ; and since along the adiabatics no heat is taken in or given out (adiabatics are, on this account, sometimes called *isentropics*), there is no change of entropy at all.

The net change of entropy in a thermal process is a criterion of the wastefulness of the process. For practical purposes, heat is useful only when it can be transformed into work. In the Carnot cycle, all the heat that is not given up

to the refrigerator is converted into work, none being wasted, and we have seen that the net change of entropy in the process is zero. But let us imagine a direct exchange of heat—whether by conduction, convection, or radiation—between bodies at different temperatures, as in the method of mixtures for determining specific heats, for example. When a small quantity of heat, dH , leaves the hot body, at temperature θ_1 , for the cold body, at temperature θ_2 , the former loses entropy $\frac{dH}{\theta_1}$ and the latter gains entropy $\frac{dH}{\theta_2}$. Since $\theta_1 > \theta_2$, the entropy lost is less than that gained, so that, on the whole, there is an increase of entropy. This applies to every increment of heat transferred during the process of equalization of temperature, so that when equilibrium is attained the entropy of the mixture is greater than before.

This gain of entropy means that heat has been wasted. It might have been converted into work, but instead it was allowed to flow spontaneously to a lower temperature, and this process is irreversible. In the Carnot cycle there is no exchange of heat between bodies at different temperatures. Here the process is reversible, there is no gain of entropy, and there is no waste. We may say, then, that if, in any process, there is a net gain of entropy, there is a waste of motive power of heat. We cannot by any process at all decrease the entropy of a whole system. We may decrease the entropy of a part, but there is always a gain of entropy somewhere else which is at least equal to the loss. This is because heat will not naturally flow against a temperature gradient. To make it do so we have to expend work, whereas the opposite process takes place spontaneously. These facts are summed up in the *Second Law of Thermodynamics*, which may be expressed by saying that heat will not of itself pass from a colder to a hotter body ; or, alternatively, that in order to diminish the entropy of a system we must increase the entropy of another system by at least an equal amount.

Molecular Interpretation of Entropy

On the molecular theory, entropy is often interpreted as the degree of disorganization of the molecular motions. In order that a body (*e.g.* a gas in a cylinder) shall do work, its molecules must act in one direction : thus they must push upwards in order to raise the piston. There is therefore a certain degree of organization in their motions, and when this is lost and the motions are at random, the gas has lost its power of doing work. Its entropy has then reached its maximum value. So long as the gas is not at the absolute zero of temperature we can always organize the motions of its molecules if we have a refrigerator at a still lower temperature, by making it the source in a Carnot engine, and so obtaining some work from its heat. The heat given up to the refrigerator, however, is then not available for transformation into work unless we have another refrigerator at a still lower temperature. Heat energy at the lowest available temperature cannot be converted into work. The random motions of its molecules cannot be converted into systematic motions.

The essential fact about the Carnot cycle, then, which makes its efficiency the greatest possible for the given temperatures of source and refrigerator, is that whenever heat is taken in or given out, the temperature is constant, and, whenever the temperature has to change, the change is made adiabatically, without passage of heat. It is only thus that we can prevent an increase of entropy, and so a waste of energy.

EXERCISES

1. State the second law of thermodynamics, and explain why the efficiency of a heat engine is greatest when the processes which take place in it are reversible.
2. Describe the thermodynamic scale of temperature, and show that it implies the existence of an absolute zero.

3. The efficiency of a Carnot cycle working between two temperatures, θ_1 and θ_2 , is $\frac{1}{8}$. If the lower temperature were changed to $\theta_3 = \theta_2 - 70^\circ$, the efficiency would be doubled. Find the values of θ_1 , θ_2 , and θ_3 .
4. Explain what is meant by *entropy*, and show that when heat passes directly between bodies at different temperatures, the total entropy is always increased.

CHAPTER XI

CHANGE OF STATE

IN Chapter IX. we considered the general theory of gases when no change of state was supposed to occur and the only forms of energy involved were heat and work. We must now consider the change of state from solid to liquid and liquid to gas. In these cases, heat or work may be converted into latent heat, and *vice versa*.

SOLID/LIQUID CHANGE

Melting-Point

Most, if not all, solids can be changed into liquids by supplying heat, and liquids can be changed into solids by subtracting heat. The temperature at which the change takes place is called the *melting-point* of the solid, or *freezing-point* of the liquid. It is sometimes a perfectly definite temperature, as when ice changes into water ; and sometimes rather indefinite, as when wax gradually softens. The melting-point can, however, be changed by various influences, *e.g.* (a) dissolved substances, (b) change of pressure, (c) absence of nuclei. Let us consider these in turn.

(a) The freezing-point of a liquid is lowered by the presence of substances in solution. This is exemplified by the common practice of adding salt to ice on the pavements on cold days. The melting-point is lowered and the ice is changed to the liquid state. A mixture of ice and salt is a very useful "freezing mixture" : it can exist as a liquid well below 0° C., the normal freezing-point of water.

(b) The effect of pressure on the melting-point of a solid is small. Solids which, like ice, contract on melting, have their melting-points lowered, and those which expand on melting have their melting-points raised, by increased pressure. An increase of one atmosphere lowers the melting-point of ice by about 0.0074° C. The smallness of this quantity explains why it is unimportant to consider ordinary barometric changes when marking the lower fixed point of a thermometer.

This effect, small though it is, has interesting consequences. It accounts for the flowing of glaciers, for instance : the weight of the glacier lowers the melting-point of the lower surface, which therefore becomes liquid, on which the remainder can float. Again, a wire, hung over a block of ice at 0° C. and weighted at each end, cuts through the ice, which unites again into a single block after the wire has passed through. The melting-point is lowered where the wire presses and the wire sinks into the liquid, but as soon as the pressure is released the water above the wire freezes into ice again.

(c) It appears that a liquid will not solidify when cooled to its freezing-point unless a nucleus of some kind—such as a dust particle, or an irregularity on the surface of the containing vessel—is present on which solidification can take place. A liquid at a temperature below its freezing-point is said to be *super-cooled*. Water can thus exist several degrees below 0° C. , but if for any reason solidification begins, it rapidly extends and the temperature rises at once to 0° C.

An important example of this occurs in the upper atmosphere. At temperatures a few degrees below zero, and in the absence of dust, water drops remain liquid. When they strike an aeroplane wing which is passing through them, they immediately solidify and adhere to the wing, causing the very troublesome “icing” well known to airmen. Had the drops been solid they would have rebounded from the wing without causing any trouble.

LIQUID/VAPOUR CHANGE

Boiling Point

There is one important likeness and one important difference between the change from liquid to vapour and that from solid to liquid. The likeness is that, just as when a solid is heated there comes a point (the melting-point) at which the temperature remains stationary until liquefaction is complete, so when a liquid is heated there comes a point (the *boiling-point*) at which the temperature remains stationary until vaporization is complete. The difference is that whereas the liquid state does not exist below the melting-point (apart from super-cooling) the vapour state exists at any temperature. Thus, on a cold day, when the temperature is well below 0° C., ice is dry ; it has no water on its surface. When the temperature is above 0° C., however, the air contains water vapour though it is far below the boiling-point of water ; and, on a dry day, water spilt on the ground "dries up"—i.e. changes into vapour without approaching the boiling-point. Indeed, there is water vapour, or "ice vapour," in the air even on the coldest day, for solids to a small extent pass directly into vapour without going through the liquid phase. This process—known as *sublimation*—is most obvious with solids that can be smelt : it is the vapour they give off that affects our olfactory sense.

Saturation Vapour Pressure

If a quantity of liquid is contained in a closed space, some of it vaporizes until the pressure reaches a certain value which is perfectly definite for a particular liquid at a particular temperature. Fig. 40 represents four barometer tubes—*a*, *b*, *c*, and *d*. *a* is an ordinary barometer containing a vacuum (or, strictly speaking, mercury vapour) above the mercury surface. In *b*, *c*, and *d* small quantities of water, alcohol, and ether, respectively, exist above the mercury.

The differences in height of the mercury columns are due (apart, of course, from the negligible weight of the small quantities of liquid) to the pressures of the vapours, which are thus seen to vary for different substances. Addition or subtraction of liquid makes no difference to the pressure, which is definitely fixed in each case so long as some liquid remains. That pressure is known as the *maximum* or *saturation* pressure for the temperature of the vapour.

Now let us suppose the tubes to be surrounded by a bath which can be raised to any desired temperature, and let this bath be gradually heated. The level of mercury falls in each tube, because the vapour pressure is being increased by the evaporation of liquid. When the level falls to that of the mercury in the outside dish, the pressure is, of course, equal to the pressure of the atmosphere, and at this point the liquid above the mercury boils. This happens at 34.9° C. for ether, 78.4° C. for alcohol, and, of course, 100° C. for water. The boiling-point may thus be defined as the temperature at which the saturation pressure of the liquid is equal to the external pressure.

Boiling-point and Pressure

It follows that we can change the boiling-point by changing the external pressure. This change is much greater than the corresponding change of the melting-point. For instance,

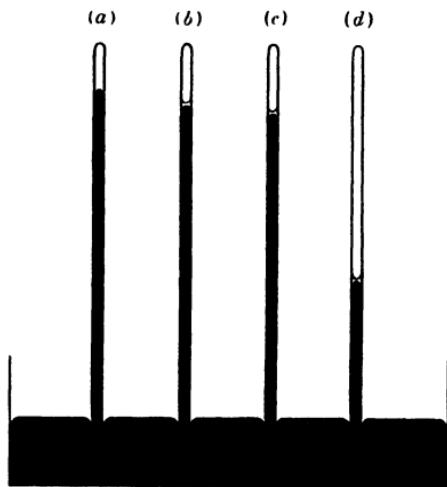


FIG. 40

Comparison of Vapour Pressures

- (a) Vacuum (b) Water
- (c) Alcohol (d) Ether

whereas an increase of pressure of one atmosphere lowers the freezing-point of water by 0.0074° C. , it raises the boiling-point more than 20° C. Similarly, a lowering of pressure lowers the boiling-point. On a high mountain water boils at considerably below 100° C.

Phenomena of Vapour Pressure

To understand clearly the phenomena of vapour pressure, suppose the cylinder in Fig. 33 contains a liquid in equilibrium with its vapour, the piston being at a certain height. The pressure below the piston will be the maximum pressure of the vapour, for, if it is less, liquid will evaporate until this pressure is reached. If the temperature is below the boiling-point, however, the external pressure will be greater than this: hence the piston will fall. If the cylinder contained only a permanent gas, the consequent decrease of volume would, of course, be accompanied by an increase of pressure, according to Boyle's law, until the internal and external pressures became equal, and the piston would come to rest at that stage. But the vapour of the liquid is already at its maximum pressure, so what happens is that as the piston falls, vapour condenses into liquid. This goes on until no vapour is left, and the piston rests on the surface of the liquid.

Now suppose we diminish the external pressure again until the saturation pressure is reached. A slight further diminution makes the internal greater than the external pressure, so that the piston begins to rise. This enlargement of the volume of the cylinder, however, does not diminish the pressure of the vapour as it would do that of an ordinary gas, for any tendency of the vapour pressure to diminish would at once be counteracted by evaporation of liquid so as to maintain the saturation pressure. The result is that the piston continues to rise until all the liquid has evaporated; the liquid, in fact, boils, since its saturation pressure has become as great as the external pressure. When no liquid is left, the piston rises through the very slight distance neces-

sary to reduce the pressure of the gas inside the cylinder to equality with the external pressure.

Saturated and Unsaturated Vapours

So long as liquid remains in the cylinder, therefore, the piston can be in equilibrium only if it is resting on the surface of the liquid, unless it so happens that the external pressure is exactly equal to the saturation pressure of the liquid, in which case the piston can rest anywhere. A vapour in contact with its own liquid thus has very different properties from a vapour existing alone ; the one is called a *saturated* vapour, and the other an *unsaturated* vapour. Ordinary gases, like air, oxygen, nitrogen, hydrogen are normally unsaturated vapours. At very low temperatures they condense, but otherwise they are never in contact with their liquid forms. The gases we have so far considered, for which Boyle's and Charles's laws approximately hold, are of this kind ; they are far from the saturation point, and the term "gas" is usually reserved for them : "vapours," by contrast, are gases near their condensation points. This distinction is rather vague, however, and a much more definite one is that between saturated and unsaturated vapours.

Dalton's Laws

If two liquids are mixed together in a closed space, each exerts its vapour pressure independently of the other, and the total pressure is the sum of the two saturation pressures. The pressure of each is then called its *partial* pressure—a general term applicable also to the pressure of an unsaturated vapour in a mixture of vapours of any kind. The partial pressure of a vapour is at least approximately the same as if the other vapours were absent. This is partly expressed by the following two laws stated by Dalton and generally known as "Dalton's Laws" :

(1) The pressure of a vapour which will saturate a given space is the same whether the space contains a gas or is a vacuum.

(2) The total pressure exerted by a mixture of two vapours, saturated or unsaturated, is the sum of the pressures which they would exert separately if each occupied the same space alone.

Humidity

The most important application of these principles is to the water vapour in the atmosphere. The approximate accuracy of Dalton's laws permits us to apply data we obtain from experiments on pure water vapour. Thus the saturation pressure of water vapour at a given temperature is the same in the air as in a barometer tube. The condition of the atmosphere as regards its water-vapour content is measured by its *humidity*, or *relative humidity* (both terms are used for the same quantity), which is defined as the ratio of the actual pressure of water vapour in the atmosphere to the saturation pressure at the same temperature. The latter quantity is found from tables which record the results of experiments similar to those already described with the barometer tubes. These experiments were made by Regnault with considerable accuracy. It is not easy to determine directly the *actual* pressure of water vapour in the atmosphere. It is simpler to cool a sample of air until a temperature is reached at which that pressure is the *saturation* pressure. The tables just mentioned then tell us at once what that pressure is.

The temperature at which the actual pressure of water vapour is the saturation pressure is called the *dew point*. It is determined by instruments known as *hygrometers*, of which there are several types.

The Vapour Pressure of Ice

It has been mentioned that solids can pass directly into vapour—a process known as *sublimation*. Since, as we have seen, water can be super-cooled, we can have, at a temperature a little below 0° C., ice and vapour in equilibrium, or water and vapour in equilibrium. It is found that the saturation

pressure of the vapour over ice is slightly less than that over water at the same temperature, as is shown by Fig. 41, in which the abscissae represent temperatures and the ordinates the saturation pressures. Below 0° C. the dotted line represents the vapour pressure of water, and the continuous line that of ice.

This means that when the vapour is cooled below 0° C., ice has a greater tendency to form than water. For, if we take a quantity of vapour in a state represented by a point between the dotted and continuous lines, its pressure is above the saturation pressure for ice, but below that for water. Hence, if possible, ice will form.

Vapour Pressure over Curved Surface

Another fact concerning vapour pressure which is not without importance in meteorology is the change of vapour pressure with curvature of the liquid surface. This follows from the fact that a liquid, in the usual case, rises in a capillary tube (see Fig. 23, p. 69). For, when in equilibrium, the vapour pressure above the concave surface of the liquid in the capillary tube is the same as that at the same level outside, which is less than that at the outside surface of the liquid by the weight of a column of vapour with a height equal to that of the liquid in the tube. On the other hand, when the surface is convex, as in the case of a small water drop, the vapour pressure is *greater* than that over a plane surface. This may be regarded as a consequence of the tendency of the surface to be as small as possible (see p. 64), for evaporation from a plane surface does not change the

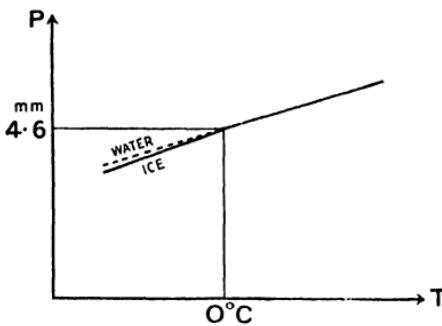


FIG. 41

Vapour pressures of Ice and Water near 0° C.

surface area, whereas evaporation from a small drop diminishes the surface. There is an additional tendency, therefore, for a small drop to evaporate, and its vapour pressure is, consequently, increased.

This fact may be connected with the need for nuclei on which condensation can take place, for, like freezing, condensation from vapour to liquid will not occur in the absence of nuclei (see p. 165). For if a very small water drop were formed, the curvature of its surface would be very great, and its maximum vapour pressure therefore increased. Hence it would tend to evaporate and disappear. A nucleus, however, would necessarily give the drop an initial size larger than the volume of the water itself, and hence a smaller curvature. The saturation pressure would therefore not be so great, and the drop could persist and grow.

Latent Heat and Total Heat

The change from liquid to vapour, like that from solid to liquid, requires the conversion of heat into latent heat. The latent heat of vaporization is defined quantitatively as the amount of heat required to change 1 gram of liquid into vapour without change of temperature. We have seen, however, that this change may occur at various temperatures, and we must inquire whether the latent heat depends on the temperature at which the change takes place. It is found that it does, and that, as we should expect, the higher the temperature the smaller is the latent heat. Regnault gave the following formula for the latent heat of vaporization of water, L , at temperature t ° C. :

$$L_t = 606.5 - 0.695t \dots \dots \quad (11.1)$$

The latent heat at 0° C. is thus 606.5 calories, and the value at the normal boiling-point, 100° C., is 537 calories.

This is a large amount of heat. More than five times as much heat is required merely to evaporate water when it

has reached 100° C. than that expended in raising it from 0° C. to the boiling-point. We experience the loss of heat when water dries on our skin by a feeling of coldness ; the latent heat is obtained from the heat of our bodies. Hence it is preferable when wet either to wipe off the moisture or to wrap up tightly so as to prevent evaporation, unless the latent heat can be supplied from an external source such as the Sun or a fire. The cooling effect of perspiration is due to a similar cause.

A quantity important in engineering is the *total heat* of a substance, which was defined by Regnault as the amount of heat required to raise unit mass of the substance from 0° C. to a temperature t° C., and then to evaporate it at that temperature. The total heat obviously varies with temperature. For water, the first part is clearly t calories, so that Q_t , the total heat of steam at t° C., is given by

$$Q_t = t + L_t = 606.5 + 0.305t \dots \quad (11.2)$$

Factors affecting the Liquid/Vapour Change

The change from liquid to vapour is affected by the same factors as that from solid to liquid, namely, dissolved substances, change of pressure, and absence of nuclei. The presence of dissolved substances raises the boiling-point. Since it also lowers the freezing-point, we see that it extends the range of liquidity of a substance in both directions. At a given temperature below the boiling-point the saturation vapour pressure of a solution is less than that of the pure solvent. This is consistent with the effect of solution on the boiling-point, for when the normal boiling-point is reached, the vapour pressure of the solution is less than the external pressure, and a further rise of temperature is therefore needed to make the liquid boil.

The effect of pressure on vaporization we have already considered at some length. The necessity for a nucleus is just as stringent in condensation from vapour to liquid as it

is in freezing. A space can become *supersaturated* (*i.e.* the vapour pressure can rise above the saturation pressure for the existing temperature) if there are no nuclei on which condensation can occur, but as soon as condensation begins, the pressure falls immediately to the saturation value. This fact has been applied to the counting of the number of dust particles in a sample of air. The air is saturated with water vapour and then suddenly cooled by expansion. Condensation takes place only on the dust particles, which, loaded by water drops, fall to the bottom of the chamber, where they can be counted.

A closely related matter is the necessity for bubbles or cavities in a liquid in order that it shall boil. Evaporation from the air-surface of a liquid in a vessel takes place at all temperatures ; the characteristic of boiling is that evaporation takes place from the interior of the liquid also, and for this small spaces, such as bubbles, are necessary into which the vapour can pass and rise to the surface. When the vessel is very clean and the water contains no air, the surface evaporation alone is not rapid enough to allow all the heat supplied to be transformed into latent heat, and the temperature of the liquid consequently rises above the normal boiling-point. When at last boiling does begin, it does so with violence, and the temperature quickly drops to the boiling-point.

Critical Temperature

When a gas is cooled sufficiently it changes into a liquid. The process is facilitated if the pressure is increased also. This is exemplified by the condensation of steam. At 1 atmosphere pressure, steam must be cooled to 100° C. before it condenses, whereas at 15 atmospheres it condenses at a little below 200° C. 100° C. and nearly 200° C. are, of course, the boiling-points at these pressures. If we consider the liquefaction of one of the "permanent" gases, such as oxygen, in the light of these facts, we see that we may expect

the gas to liquefy at a higher temperature if we increase the pressure on it.

The question therefore arises : can we compress oxygen so much that it liquefies without the temperature being changed at all ? The answer is that it depends on what the temperature happens to be. If it is above a certain value, known as the *critical temperature*, then no amount of increase of pressure will turn it into a liquid. If it is below this temperature, however, increase of pressure will liquefy it. (The same thing is true of steam, but here the critical temperature is very high. It is true, in fact, of any gas or vapour.)

Liquid/Vapour Isothermals

This may be understood better if we look at an indicator diagram showing the isothermals for a vapour below its critical temperature.

Let us start (Fig. 42) at A_1 , with 1 gm. of unsaturated vapour, and gradually increase the pressure. So long as the vapour remains unsaturated, the isothermal will follow the usual course for a gas, namely, A_1B_1 ; but when the pressure reaches the saturation value at B_1 , condensation begins, and the pressure then remains constant until all the vapour is condensed.

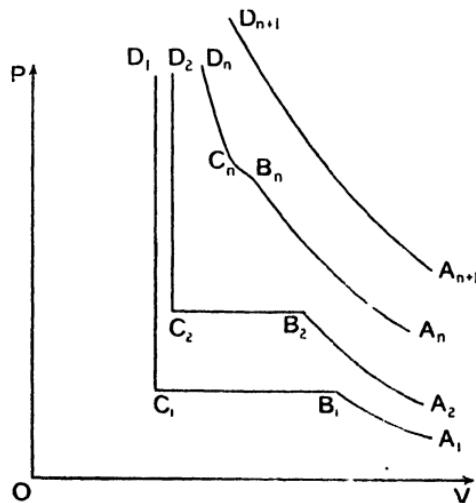


FIG. 42

Isothermals for Liquid and Vapour

This is represented by the horizontal line, B_1C_1 . At C_1 we have liquid only, and then an increase of pressure changes the volume very little, so the curve continues as an almost vertical

line, C_1D_1 . The complete isothermal, then, has the form $A_1B_1C_1D_1$.

Now let us repeat the process at a higher temperature. We then obtain the curve $A_2B_2C_2D_2$. It can be shown that B_2 lies to the left of B_1 , and C_2 clearly lies to the right of C_1 , since the liquid expands with rise of temperature. The "liquid" portion of the curve is therefore shortened at both ends. On further increase of temperature the same tendency is continued, and ultimately we arrive at a temperature at which B_nC_n shrinks to a mere kink in the curve. This temperature is the critical temperature. If the temperature is higher, the isothermal takes its ordinary form for a gas, and approximately follows Boyle's law.

Elaborate experiments on the isothermals of carbon dioxide were carried out by Andrews. He gradually increased the pressure on a sample of the gas, and noted the successive volumes. This procedure was repeated at various temperatures, and the isothermals were drawn accordingly. Andrews found that the critical temperature for CO_2 was $30.92^\circ C$. At this temperature the pressure at which the substance was on the point of condensation (the *critical pressure*) was about 73–75 atmospheres, and the volume of unit mass (the *critical volume*) was about $\frac{1}{16.9}$ of the volume at 1 atmosphere pressure; Boyle's law, of course, would give a volume of from $\frac{1}{7.3}$ to $\frac{1}{7.5}$ of the volume at 1 atmosphere, so that vapours near the condensation point depart considerably from this law. At the critical point the substance is in a sort of intermediate condition between liquid and vapour. It has almost the same volume as that of the liquid in the neighbouring liquid state, but the molecules have the freedom of movement of those of a gas. The surface of separation between liquid and vapour disappears.

Liquefaction of Gases

It is clear from these facts that if we are to liquefy one of the "permanent" gases, we must first cool it below the critical

temperature. The most satisfactory method of doing this is to make use of the slight cooling which occurs on expansion through the performance of internal work (p. 142). The gas is forced through a spiral tube, at the end of which it escapes suddenly into a region of lower pressure, a slight cooling thus taking place. The cooler gas then returns round the outside of the tube, and is again forced along the same path. In returning it slightly cools the oncoming gas in the tube, which therefore escapes at a lower temperature and is cooled a little more. This, in its turn, still further cools the gas following it, so that a continuous process of cooling goes on as the gas goes round and round the circuit. This continues until at last the gas is liquefied and issues as drops which can be collected.

This "regenerative" process, as it is called, is used on the commercial scale for gases such as air. Hydrogen and helium require specially low temperatures, and methods into which we cannot enter here are used to liquefy and even solidify them. By modern methods it has been possible to reach temperatures within a small fraction of a degree above the absolute zero.

EXERCISES

1. State the chief physical conditions by which the melting-point of a solid is influenced, and give examples, occurring in nature, of their effect.
2. Explain what is meant by the *saturation pressure* of the vapour of a liquid, and discuss its relation to the boiling-point.
3. Describe the difference in behaviour of saturated and unsaturated vapours when subjected to changes of pressure and temperature.
4. The saturation vapour pressure of water vapour is 12.78 mm. at $15^{\circ}\text{C}.$, and 8.04 mm. at $8^{\circ}\text{C}.$ Assuming that between these temperatures the pressure varies linearly with

the temperature, find the humidity when the temperature of the air is 14° C. and the dew point is 9° C.

5. Explain the term "critical temperature." Draw a series of isothermals for a substance over a range of temperature including the critical temperature, and explain how the state of the substance varies along the curves.

6. Give an account of the phenomena of change of state in terms of the molecular theory of matter, explaining in particular the transformation of heat into latent heat, or *vice versa*. Why does the latent heat of vaporization of a liquid decrease with rise of temperature?

7. The volume of a mixture of air and saturated vapour at constant temperature is halved, and it is then found that the total pressure is increased by one-third. Find the ratio of the original pressures of air and vapour.

CHAPTER XII

THE TRANSFER OF HEAT

WE consider now the three processes of transfer of heat—conduction, convection, and radiation—described in an earlier chapter.

CONDUCTION

Diffusivity

The difference in thermal conductivity between bodies is easily illustrated by holding a piece of wire and a piece of wood in a flame. The wire soon becomes unbearably hot, whereas the wood can be held until the flame reaches the fingers.

Although it is true that the metal has the greater conductivity for heat, this is, however, not necessarily proved by such an experiment, for what we feel is the rise of *temperature*, not the accumulation of heat. A given quantity of heat raises bodies to temperatures depending on their heat capacities, and what our experiment shows is that the conductivity for *temperature*, to speak metaphorically—the *diffusivity*, as it is sometimes called, in spite of the possibility of confusion with the quantity defined in Chapter V—is greater for the metal than for the wood. Diffusivity is defined quantitatively as the thermal conductivity divided by the heat capacity per cc.

Thermal Conductivity

To see the factors on which thermal conductivity depends, let us imagine a slab of material (Fig. 43), of which one side is kept permanently at a temperature θ_2 , and the other side

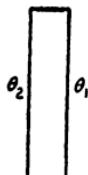


FIG. 43
Thermal
Con-
ductivity

permanently at a lower temperature, \$\theta_1\$. Heat will then flow from the hotter to the colder side, and since everything is kept in a steady state, the rate of flow of heat must be constant. The amount, \$H\$, which flows is then clearly proportional to the time of flowing, \$t\$, and to the area, \$A\$, of the faces of the slab, so that

$$H \propto At \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (12 \cdot 1)$$

It depends also on the difference of temperature, \$\theta_2 - \theta_1\$, and on the distance, \$l\$, between the faces—the thickness of the slab—but the character of the dependence on these factors is not obvious. We may take it as a fact of experiment that, provided the thickness and temperature difference are not very great, the heat is directly proportional to \$\theta_2 - \theta_1\$ and inversely proportional to \$l\$. Hence

$$H \propto \frac{At(\theta_2 - \theta_1)}{l} \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (12 \cdot 2)$$

$$\text{or } H = K \frac{At(\theta_2 - \theta_1)}{l} \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (12 \cdot 3)$$

where \$K\$ is a constant. It is called the *thermal conductivity* of the substance of the slab. Expressed in words, it is the quantity of heat conducted per second normally across 1 sq. cm. of a substance in the direction of a temperature gradient of \$1^\circ\$ C. per cm. A more exact form of equation (12.3) is

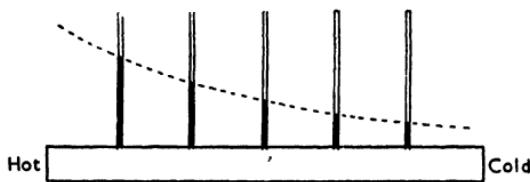
$$H = KAt \frac{d\theta}{dl} \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad (12 \cdot 4)$$

Measurement of Thermal Conductivity

Determinations of the conductivity of materials depend on the measurement of the various quantities in this equation. The most difficult of these to measure is \$H\$, for we must eliminate all heat that is not actually conducted through the

material. One satisfactory method for solids is that of Forbes, in which the two ends of a bar were kept at constant but different temperatures. Thermometers at various points along the bar gave the distribution of temperature (Fig. 44), and when their readings remained constant, the flow of heat was steady. At any chosen cross-section, A could be determined and $\frac{dl}{d\theta}$ was given by the slope of the curve through the thermometer levels. It remains to determine $\frac{H}{t}$, the heat flowing across the section per second. Since the conditions were steady, all the heat passing the section must have been radiated from the surface of the bar beyond the section;

FIG. 44
Measurement
of Thermal
Conductivity



otherwise heat would be accumulating somewhere, and the temperature there would rise.

To find this quantity, a subsidiary experiment was made, in which a bar of the same material and surface was heated, and its rate of cooling (*i.e.* of fall of temperature) observed at various temperatures. Denoting this by $\frac{d\theta}{dt}$, the rate of heat

loss through the cooling $\left(\frac{dH}{dt}\right)$ was $MS \frac{d\theta}{dt}$, where MS denotes the heat capacity of the bar. Returning now to the first bar, we know θ at all points beyond the section, and therefore we can determine $\frac{d\theta}{dt}$ for all such points from the subsidiary experiment. The rate of heat loss is therefore known from the heat capacity of the bar.

Special methods have to be used for liquids and gases, whose conductivities are much smaller than those of solids.

This is intelligible in terms of the molecular theory, since conductivity is conceived as the passing on of energy from molecule to molecule by collision. When the molecules have some freedom of movement, as in liquids, and collisions are rare, as in gases, we should expect less direct transfer of kinetic energy than we get in solids, where the relative positions of the molecules remain the same. The very widely varying conductivity of different solids is not so easy to understand. It is connected with the presence of free electric charges in substances. Thus, metals, which are good conductors of electricity, are good conductors of heat also, while electrical insulators are bad conductors of heat.

CONVECTION

Convection is the commonest process by which heat is transferred from one part of a fluid, whether liquid or gas, to another. It is exemplified in the boiling of kettles of water by heating them at the bottom, in the systems of domestic heating by water circulation, and in many other common phenomena. Convective circulation in gases can be seen in the shimmering effect over a heated surface, as, for instance, when one looks at objects through the air just over a flame. The heated air rises by convection, and since the refractive index of air depends on its density, we get rapid fluctuations in the direction in which a particular object is seen.

Convective Equilibrium

As we rise in the atmosphere, the temperature at first falls. The air, therefore, resembles a fluid heated at the bottom, and a circulation tends to be set up. This is opposed, however, by the fact that the density of the upper layers is less than that of the lower through causes independent of temperature, so that we have two opposing tendencies. The state of equilibrium—*convective equilibrium*, as it is called—is that in which

a mass of air moving upwards would be cooled by its expansion to the same temperature as that already existing in its new situation. In that case there would be no tendency for air to ascend or descend, and it would remain quiescent.

In this state there is no exchange of heat between air which might rise, and that into which it moves, since they are both at the same temperature. The relation between pressure and temperature at different heights is therefore given by the adiabatic equation (9.22). This gives a fall of temperature as we go upwards of about 1° C. for every 100 metres.

RADIATION

Radiant Energy

Transfer of heat by radiation differs from the other modes of transfer in that it involves the transformation of heat into something else which is later transformed back into heat again. For instance, the heat we receive from the Sun passes through 93 million miles of space which is almost empty of matter. There is no possibility of molecular kinetic energy passing directly from the Sun to us, so we must suppose that the heat energy of the Sun changes into something which crosses the intervening space and is re-converted into heat when it meets matter on the Earth. The same thing occurs on a smaller scale when a fire warms our bodies although the air between remains cold.

We have already (pp. 102-3) spoken of the ether—the medium which is assumed to make possible the conveyance of energy by wave motion across space empty of matter. The motions of the molecules of a hot body originate waves in the ether, which travel out in all directions until they encounter matter. They may then be reflected from, or transmitted through, the matter, or they may be absorbed by it. In the last-named case the waves cease to exist, and their energy is converted into molecular motions in the receiving body. That is what happens when heat is radiated from one body to another.

Prévost's Theory of Exchanges

Experience shows that radiation takes place from a hotter to a colder body. If two bodies at different temperatures are placed near one another, then (if radiation to and from other neighbouring bodies be prevented or allowed for) heat passes by radiation from the hotter to the colder body until the temperatures are equal. We find it convenient, however, to express this fact by supposing that all bodies, at whatever temperature they may be (assumed above the absolute zero), are radiating heat into space, and that the rate of radiation for each body increases as its temperature rises. Each of our two bodies, then, is radiating, but the colder receives more than it emits, and the hotter emits more than it receives. The former therefore rises and the latter falls in temperature, until the temperatures are equal, when each receives heat at the same rate as it radiates, and no further change takes place. This way of looking at the matter is known as *Prévost's Theory of Exchanges*.

Surface Radiation

We find by experiment that the rate of radiation of a body depends on the temperature, and also on the nature of its surface, and on nothing else ; the interior of the body makes no difference. (The heat capacity of the body, of course, will determine the rate at which its *temperature* falls for a given rate of radiation of heat, but that is another matter.) Since, then, we can change the rate of radiation of a body at a given temperature by changing its surface, it would seem that we could make heat pass between bodies at the same temperature, thus allowing one spontaneously to get colder and the other hotter. For suppose we have two similar bodies at the same temperature, so that each receives from the other as much heat per second as it radiates thereto. Cover one with a better radiator, keeping the temperature constant. It will then radiate more than before, while the same amount falls on it.

Its temperature should therefore fall, while that of the other, which is now receiving more radiation, should rise.

This, however, does not happen; no change of temperature takes place. We conclude, therefore, that the better radiating surface must be also a better *absorbing* surface, so that it can accept as well as radiate more heat than it did before. In this way we arrive at the general rule that *the emissive and absorptive powers are equal to one another for every kind of surface.*

Black Body Radiation

Obviously, if a surface absorbs all the radiation which falls on it (*i.e.* if all such radiation is converted into molecular energy, none being transmitted and none reflected), it has the maximum possible absorptive power. It must therefore have the maximum possible emissive power also. Such a body is called a *black body*—a rather unsatisfactory name, derived from the fact that the bodies which approximate most closely to this state are dull black, *e.g.* lampblack. The *emissive power* of a body is defined as the ratio of its emission to that of a black body at the same temperature; and the *absorbing power* as the ratio of its absorption to that of a black body. Since a black body absorbs *all* radiation falling on it, an alternative definition of the absorbing power is the fraction of the incident radiation which it absorbs.

Experimental Study of Radiation

The emissive powers of bodies are conveniently studied by means of a “Leslie’s cube”—a hollow cube filled with water which can be heated, the sides of the cube being coated with the various materials under examination. Receiving instruments, placed opposite each face in turn, indicate the rate of radiation by the rate at which they rise in temperature. The receiving instruments are usually coated with lampblack, so that they absorb as much as possible of the radiation falling on them. Instruments commonly used for this purpose are

the *thermopile* and the *bolometer* ; they will be described in the second volume of this book.

Thermal Insulation

It is fairly obvious that a good absorber of heat must be a bad reflector, for if most of the radiant energy is absorbed, there must be little left to be reflected. Similarly, good reflectors are bad absorbers, and therefore bad radiators. For this reason brightly polished kettles are more economical than sooty ones, for, being good reflectors, they are bad radiators, so that the heat which goes into them is mostly available for heating the contents. This principle underlies the use of the *Dewar flask*—the ordinary “thermos” flask. This consists of a double-walled vessel (Fig. 45), of which the walls are brightly silvered and the enclosed space is evacuated. A body in the flask therefore loses and receives very little heat, for the vacuum prevents conduction between the outside and inside, and the bright surface allows very little radiation. The temperature inside may therefore remain very different from that outside for a long time.

Radiation and Temperature

For a given surface the rate of radiation rises rapidly with increase of temperature ; it is, in fact, proportional to the fourth power of the absolute temperature for a black body. Thus, if R denotes the energy radiated per second per sq. cm., we have

$$R = a\theta^4 \quad \dots \quad \dots \quad \dots \quad \dots \quad (12.5)$$

where θ is the absolute temperature and a is a constant. If the surroundings are at temperature θ_1 , and they radiate also as a black body, the surface loses heat at a net rate of $a(\theta^4 - \theta_1^4)$. This is known as the *Stefan-Boltzmann law*.

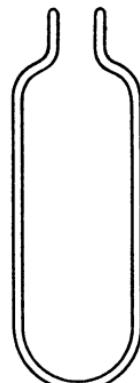


FIG. 45

Dewar
Vacuum
Flask

EXERCISES

1. Distinguish between conductivity for heat and conductivity for temperature, and state on what factors the conductivity for heat depends and how it depends on them. What additional factor is concerned in the conductivity for temperature?
2. Opposite faces of a cube of marble, whose thermal conductivity is 0.005, are kept in contact with steam and a mixture of ice and water respectively, and the remaining faces are covered with some non-conducting and non-radiating material. If the side of the cube is 10 cm. long, calculate how much ice is melted per minute. (Latent heat of fusion of ice = 80 calories.)
3. Explain Prévost's theory of exchanges, and describe how the exchanges of heat by radiation between a body and its surroundings vary as the body is raised from a temperature below to a temperature above that of the surroundings.
4. Prove that the thermal absorptive and emissive powers of a body are equal.

PART IV
VIBRATIONS AND SOUND

CHAPTER XIII

WAVE MOTION

Nature of Sound

THE sensation of sound is caused by the impact on our ears of waves transmitted through the air. These waves originate in the source of the sound, and travel through the matter (possibly solid, liquid, or gaseous) which lies between it and the hearer, ending, in normal circumstances, as vibrations in the air surrounding the ears.

That a material medium of some kind, reaching from the source of sound to the observer, is necessary can be shown most simply by enclosing an electric bell in a vessel which can be evacuated. Set the bell ringing, and then gradually pump the air out of the vessel. The sound gets fainter, and finally dies away altogether, although the bell can be seen to be vibrating as before. On readmitting air to the vessel the sound is restored. Clearly, then, sound cannot travel through a vacuum, and, moreover, since the bell is entirely enclosed within the vessel, the sound, when it is heard, does not travel only through air but through the solid material of the vessel also.

In order to understand the phenomena of sound, it is necessary to know something about the characteristics of waves in matter. We therefore begin with a general discussion of wave motion, which is the means by which a disturbance at some point in an elastic body is transmitted through the body.

Wave-Motion

Suppose we have a rod of some elastic material, and we create a disturbance at one end by a slight displacement of the molecules. This may be done in two ways (or by some process intermediate between the two, but such a case can always be resolved into a combination of those we are considering) : the molecules can be displaced along the direction of the rod by a slight compression or stretching, or they can be displaced in a direction perpendicular to the rod. The first displacement gives rise to what are called *longitudinal* waves, and the second to what are called *transverse* waves. We shall consider them in turn, but we may say at once that sound waves are of the longitudinal type, while light waves and waves on the surface of water are transverse.

Suppose a slight momentary pressure is applied to the end of the rod, so that the molecules there move closer to their neighbours, in obedience to Hooke's law. Since the elastic limit has not been passed, they return when the pressure is released, but the pressure they have exerted on neighbouring molecules has caused those molecules to move forward against the next layer of molecules, and so on along the rod. We thus see that the molecules in each successive cross-section are moved forward and then return when the pressure is released. This transmission of pressure along the rod, of course, takes time, so that when the molecules at the end first disturbed are returning to their original positions, those at a certain distance along the rod are only just beginning to feel the compression, and those still farther along have not been disturbed at all. An impulse, consisting of a compression followed by a rarefaction, is thus made to travel along the rod from one end to the other.

Now suppose that instead of a single compression we have a series of compressions, following one another at regular intervals. We might imagine, for instance, a series of hammer taps given to the end of the rod. In that case we shall have a

whole series of impulses transmitted along the rod, following one another in regular succession, and if we could see the arrangement of the molecules we should see them alternately condensed and rarefied at equal intervals all along the rod.

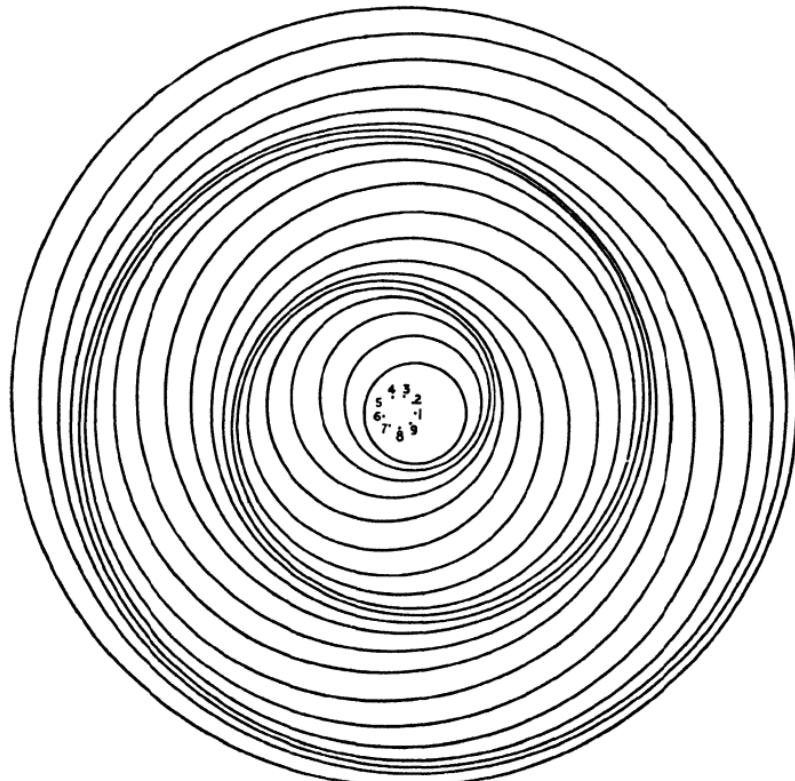


FIG. 46
Crová's Disc

Each region of condensation or rarefaction would appear to travel uniformly from the disturbed end to the other. The whole arrangement would constitute what is called a *wave-train*.

We may form a very good picture of the appearance of
(276)

such a wave-train by means of a diagram known as "Crova's disc." To make such a disc, draw on a card a small circle, and mark a number of equidistant points around its circumference, such as the points 1, 2, . . . 9 in Fig. 46. With 1 as centre, draw a circle of radius r , say, which is greater than the radius of the original circle. Now, with 2 as centre, draw another circle of radius $r + a$, where a is a little greater than the distance between the points 1 and 2. Similarly, with 3, 4 . . . as centres, draw circles of radii $r + 2a$, $r + 3a$, . . . , and so on, until the card is filled.

Now rotate the card about its centre, and, to reproduce the appearance of the vibrating rod, place in front of it another card from which a horizontal slot has been cut out so as to expose only a long narrow strip of the Crova disc extending from the original small circle to the outermost circle of the disc in one direction. The appearance of the successive condensations and rarefactions travelling from one end of the slot to the other is then vividly and accurately presented.

It is important to understand that what travels from one end of the rod to the other is not the material of the rod but only the appearance of alternate condensation and rarefaction: it is that which constitutes the wave. Any particular molecule moves only a short distance and back again time after time; it oscillates along the direction of the rod, about its fixed mean position. Similarly the circles in the Crova disc do not travel along the slot.

It should be mentioned also that this regular motion of the molecules does not constitute heat and raise the temperature of the rod. Heat consists of *random* motions of the molecules, and an organized wave motion is no more a part of the heat content of the rod than is the bodily motion given it when it is carried from one place to another.

Characteristics of a Wave-train

Three quantities connected with a wave-train are of great importance; they are the *velocity*, the *wave-length*, and the

frequency. The velocity, V , of the wave-train is the distance which any particular condensation or rarefaction travels in one second. The wave-length, λ , is the distance from one condensation or rarefaction to the next. The frequency, n , is the number of condensations or rarefactions which pass a given point on the rod in one second.

There is an important relation between these quantities. Let us choose two cross-sections of the rod separated by a distance V , *i.e.* the distance which the wave travels in one second. If, at a certain instant, a particular condensation is at the first of these cross-sections, it will therefore be at the other a second later, and all the condensations which have crossed the first section in that time will lie along the interval between the two cross-sections. But this number is, by definition, n ; hence there are n waves in a distance V . But each wave has a length λ . Hence n waves have a total length $n\lambda$, and therefore

$$V = n\lambda. \dots \dots \dots \quad (13 \cdot 1)$$

This is a fundamental equation in the theory of wave motion.

It is found that the velocity of the waves is a property of the medium transmitting them, and of nothing else. It is, in fact, given by

$$V = \sqrt{\frac{E}{\rho}}. \dots \dots \quad (13 \cdot 2)$$

where E is the modulus of elasticity (Young's modulus for a thin rod and the bulk modulus for a medium extending in all directions), and ρ is the density of the medium. The velocity is the same for different wave-lengths or frequencies. The frequency, of course, can be made almost what we like, for we can choose the frequency with which we apply the hammer taps. It follows that the wave-length will vary inversely as the frequency, for their product, by equation (13.1), is a constant quantity. In a given material, then, in a given

physical condition, we cannot have waves of any frequency and also any wave-length we may wish. We can choose either of these quantities, but the other will then be determined for us.

Transverse Waves

A transverse wave occurs when the molecules at the end of the rod are displaced not along the length of the rod but at right angles thereto. The existence of such a wave can be made evident visually by stretching a flexible material, such as a rubber tube, along a table, and quickly lifting one end a short distance and releasing it. The hump thus formed may be seen to travel along the rubber to the other end. The same thing is observed when a stone is thrown into water ; a transverse wave spreads out in all directions from the point where the surface molecules of the water are depressed.

If the displacement is repeated regularly we get a wave-train, just as we do with a longitudinal wave, but instead of a succession of condensations and rarefactions, we have a succession of crests and troughs. It is these displacements that travel and constitute the wave-motion ; each molecule of the medium simply oscillates up and down without moving in the direction of the wave. This may be shown in the case of water waves by putting some light material, such as a small piece of wood, on the surface of the water ; it will be seen to rise and fall while the wave-train goes continually past it. The general characteristics of waves are the same whether the waves are transverse or longitudinal, though the physical effects of the wave-motion may show wide variations. Our definitions of velocity, frequency, and wave-length apply equally well to transverse waves if we substitute crests and troughs for condensations and rarefactions, and the same fundamental equation (13.1) holds for both types of wave. We shall not, therefore, distinguish, in what follows, between one type of wave and the other unless it becomes necessary to do so.

Equation to a Wave-train

Let us consider the displacement from its normal position of a particle in a wave-train. At any instant the displacement will vary for different particles, and for any particular particle the displacement will vary at different instants. Hence, in mathematical language, the displacement, y , of any, unspecified, particle is a function of the time, t , and of the distance, x , of the particle from the origin of the disturbance. An expression giving y in terms of t and x is called the *equation* to the wave-train.

Consider the equation

$$y = A \sin 2\pi \left(nt + \frac{x}{\lambda} \right) \dots \dots \quad (13.3)$$

where n and λ are constant. If we give x a fixed value, y is a function of a single variable, t ; and as t increases continuously, the sine oscillates to and fro between the extreme values, $+1$ and -1 . Hence y oscillates between the extreme values $+A$ and $-A$. If, in order to fix our ideas, we think of a transverse wave (though, as we have said, the same thing applies to a longitudinal wave), we see that this corresponds to the behaviour of a particular particle in the wave-train; our piece of wood floating on water, for example, moves up and down between two extreme positions. The maximum distance of displacement—represented in the equation by $\pm A$ —is called the *amplitude* of the wave.

The interval of time between two successive maximum displacements of a particle in the same direction is easily found. For, if such a displacement occurs at time t_1 , say, then, in order that y shall be equal to A , we must have

$$\sin 2\pi \left(nt_1 + \frac{x}{\lambda} \right) = 1 \dots \dots \quad (13.4)$$

$$i.e. \quad nt_1 + \frac{x}{\lambda} = N + \frac{1}{4} \dots \dots \quad (13.5)$$

where N is any integer, including zero. The next instant at which the maximum displacement in the same direction

occurs will be, say, t_2 , which is given by the next integral value of N , so that

$$nt_2 + \frac{x}{\lambda} = N + \frac{5}{4} \quad \dots \dots \quad (13.6)$$

Hence, subtracting (13.5) from (13.6)

$$n(t_2 - t_1) = 1 \quad \dots \dots \quad (13.7)$$

i.e. $t_2 - t_1$, the interval required, is $\frac{1}{n}$. This interval is called the *period* of the wave. It is the interval between two successive equal displacements of a particle in the same direction. We have taken this displacement as the maximum, but the same interval of time holds good for any displacement of the same amount in the same direction. It can be seen from this that n , in equation (13.3), is actually the frequency of the waves. For, since the period of time between two successive maximum displacements is $\frac{1}{n}$ seconds, the number of such displacements per second at the point x must be n , and this, by definition, is the frequency of the wave-train.

Now let us regard t as constant, and consider how y varies with x ; i.e. we consider the variation of displacement along the wave-train at a particular instant. This variation is represented by Fig. 47. We then have y as a function of x , and the maximum displacements in a particular direction occur at those positions for which $nt + \frac{x}{\lambda} = N + \frac{1}{4}$, as before. If x_1 and x_2 are two successive values of x for which this condition is fulfilled, then we have

$$nt + \frac{x_1}{\lambda} = N + \frac{1}{4} \quad \left. \right\} \quad \dots \dots \quad (13.8)$$

$$\text{and } nt + \frac{x_2}{\lambda} = N + \frac{5}{4} \quad \left. \right\} \quad \dots \dots \quad (13.8)$$

$$\text{so that } \frac{x_2 - x_1}{\lambda} = 1 \quad \dots \dots \quad (13.9)$$

But $x_2 - x_1$ is, by definition, the wave-length. Hence λ in equation (13.3) is actually the wave-length of the train.

Arguing backwards, then, we can say that if in equation (13.3) A represents the maximum displacement, or amplitude, of a particle, and n and λ are respectively the frequency and wave-length of the train, the equation can represent a wave-

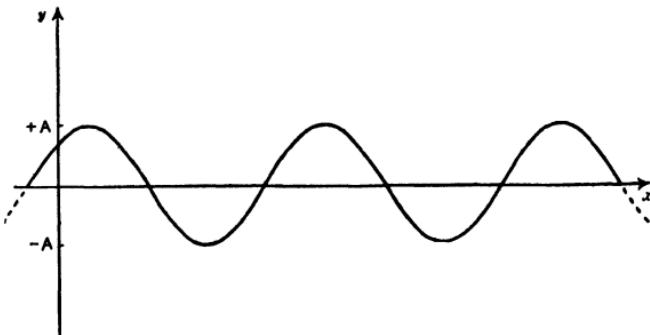


FIG. 47
Simple harmonic wave-train

train. Clearly, from our fundamental relation, $V = n\lambda$, we can write the equation as

$$y = A \sin \frac{2\pi}{\lambda} (Vt + x) \quad \dots \quad (13.10)$$

or, as yet another alternative, if we represent the period by T , we can write it as

$$y = A \sin 2\pi \left(\frac{t}{T} + \frac{x}{\lambda} \right) \quad \dots \quad (13.11)$$

Phase

The displacement is, of course, only at its maximum value at particular points at any instant: over the greater part of the wave-train it is something less than this. The amount of the displacement is usually represented by what is called the *phase*. At a distance l from the crest of a wave, the phase is given by $2\pi \frac{l}{\lambda}$ so that whenever l is an integral number of

wave-lengths the phase is an integral multiple of 2π . It is usual, in such cases, to call the phase simply 2π , whatever the multiple—thus, in effect, concentrating attention on a single wave in the train. If $l = (N + a)\lambda$, where N is an integer and a a fraction, the phase becomes $2\pi(N + a)$ and again we ignore the term $2\pi N$ and call the phase $2\pi a$. The phase of a wave may thus be anything from 0 to 2π , corresponding to the various displacements within the distance from one crest to the next.

It is clear from (13.11) that we get the same change in y whether we change $\frac{x}{\lambda}$ to $\frac{x + a\lambda}{\lambda}$, or $\frac{t}{T}$ to $\frac{t + aT}{T}$. Hence we may think of the phase as characterizing the displacement of any particle in the wave-train at some particular instant, or that of some particular particle at any instant during its vibration. What the phase represents is the amount of the displacement as a fraction of the maximum displacement, whether of a particular particle or at a particular point in the wave-train.

Fourier's Theorem

We have said that equation (13.3) and any of its alternative forms can represent a wave-train. It does not follow, however, that every wave-train will be represented thereby. We have shown that it represents the characteristics of a wave-train so far as frequency and wave-length are concerned, but not necessarily with regard to the shape of the waves, *i.e.* with regard to the way in which the displacement varies with either t or x within a particular period or between two successive crests. In other words, we have not shown that it necessarily represents the phases of the displacements correctly. It does not, in fact, always do so : there are waves of various forms which cannot all be represented by the same equation, even though they may have the same frequency and wave-length. It can be shown, however, that if the restoring force brought into play by the original displacement is proportional

to the amount of the displacement, then the oscillation will actually be represented by a sine curve, and the wave-train by equation (13.3). Now this is just what occurs with the simple disturbances we have considered, such as an elastic deformation of a rod, for, by Hooke's law, the stress is proportional to the strain. Hence equation (13.3) does actually represent the whole wave-train in such cases. A vibration of this kind is said to be *simple harmonic*.

Much more complicated disturbances, however, can occur and be transmitted through matter in wave form. In such cases the wave-train is represented by other equations. There is a very important theorem known as *Fourier's theorem*, according to which a wave-train of *any* form (satisfying certain mathematical conditions which are, in fact, satisfied in most actual cases) can be represented by a sum of expressions such as (13.3), in which the wave-lengths are respectively λ , $\frac{\lambda}{2}$, $\frac{\lambda}{3}$, $\frac{\lambda}{4}$, etc., and the frequencies n , $2n$, $3n$, $4n$, etc. The analysis of the wave-train into these components is known as *harmonic analysis*. In the general case there are an infinite number of such components, but in many actual cases a few only are necessary to represent the wave-train either exactly or with very close approximation.

Reflection of Waves

When a wave-train impinges on a rigid surface it is reflected, *i.e.* a wave-train starts from the point of incidence and travels back through the medium. That this will occur may be seen most easily in the case of longitudinal waves. Suppose, for example, one end of our rod meets a wall normally, and the impulse is given it at the other end. When the compression reaches the wall the molecules of the rod are impelled towards the wall and we have a case of impact, in which, as we know (p. 75) there will be a rebound. A wave of compression is therefore sent back through the rod, and if we have a succession of impulses we shall have two wave-trains,

one travelling forward and the other backward along the rod.

The result of this is to set up what are called *stationary waves* along the rod. The actual motion of each molecule will be given by the resultant of the displacements which the

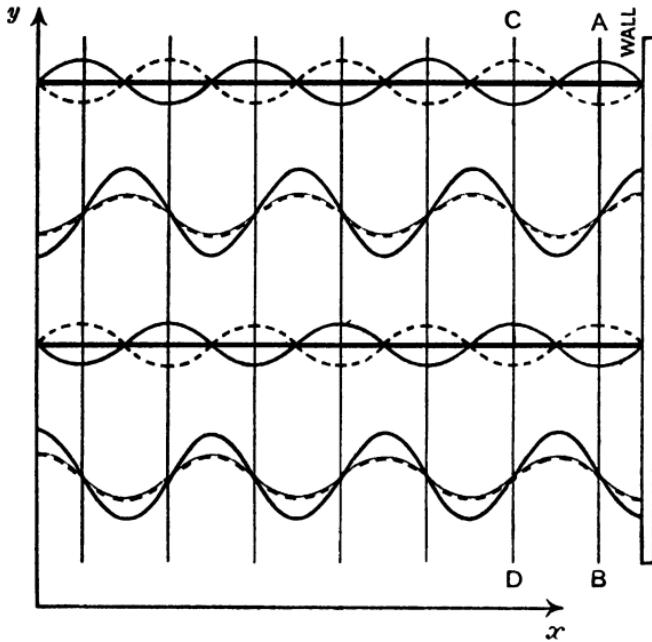


FIG. 48

Reflection of waves at a rigid wall

two waves would give it separately. Let us see what this will be.

In Fig. 48 we have represented the displacement, y , by the ordinates, and distances, x , along the rod by the abscissae. In a longitudinal wave, of course, the actual displacement is in the x direction, so that the figure is merely a graph, and not an actual picture of the motion. The thin continuous line in each diagram represents the incident wave moving from left to right, and the dotted line the reflected wave, moving from

right to left. The resultant, represented by the thick line (which gives, of course, the actual displacement) is obtained by adding the two ordinates. In the top figure the resultant is always zero, so that at the instant at which the waves have this form there is no displacement of any of the molecules. The second figure shows the state of affairs a quarter of a period later. Here the incident wave has moved a quarter of a wave-length to the right, and the reflected wave a quarter of a wave-length to the left, and we see that the two waves now coincide. The resultant displacement is therefore double that given by either. For the third figure another quarter period has elapsed, and now the two waves are interchanged with respect to the first configuration, so that again we have a resultant with no displacement. In the fourth figure we have the appearance after a further quarter period. After this the same series of changes is repeated over and over again.

Now we see from these diagrams that the molecules in the vertical lines are always undisplaced, while those half-way between them oscillate to and fro with double the amplitude of the incident wave. There is thus no appearance here of a forward or backward movement of the waves, but simply a continual up and down movement of some particles, with varying amplitudes, while others, at regular intervals, remain stationary. It is for this reason that the resultant wave is called a *stationary* wave. The points which remain always at rest are called *nodes*, and the points midway between them, which undergo the maximum displacement, are called *anti-nodes* or *loops*.

In the case just considered we have supposed the wall to be rigid, so that it resists the incident compression. If it yields—in particular, if there is no wall at all, but simply air—then it can be shown that we still have reflection, but of a different type, for there is a sudden change of phase equal to π , and the second diagram in Fig. 48, for example, has the appearance shown in Fig. 49 instead of that of two coincident

waves. The other diagrams also are changed to a form obtainable by moving the dotted curve half a wave-length to the right while keeping the continuous curve unchanged. (Actually, of course, no wall is perfectly rigid, for all solids have a finite modulus of elasticity. The difference between the two cases we are considering is more strictly stated by saying that in the former case the modulus of the wall was large, and in the latter case it is small, compared with that of the rod.)

We may see the reason for this phase change if we again

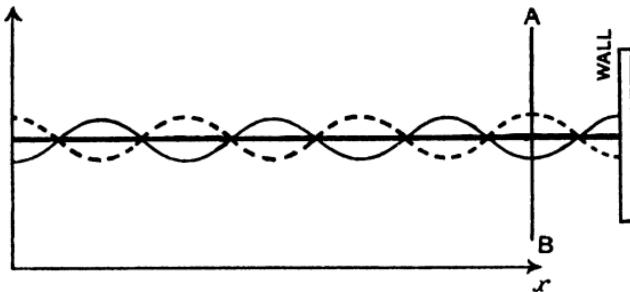


FIG. 49

Reflection of waves at a yielding wall

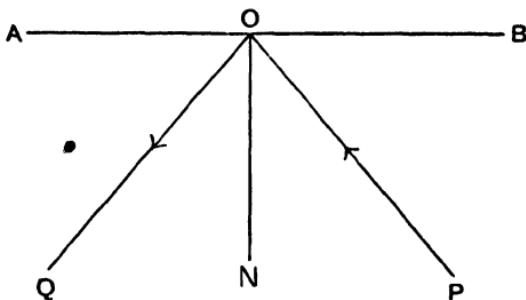
consider a compression reaching the free end of the rod. It does not now meet a resisting wall, but a region of little or no resistance, so that it suddenly expends itself into a rarefaction. The reflected wave is therefore like the incident wave suddenly advanced half a wave-length. Putting $a = \frac{1}{2}$, in the expression for the phase, we see that this is equivalent to a change of phase of π . Nevertheless, if we repeat all the diagrams in Fig. 48 with the modification described above, we see that we still have stationary waves of the same character as before, but with the nodes and loops interchanged.

We have here taken the case of a rod meeting a reflecting surface normally. If we have waves travelling in a certain direction in a continuous medium such as air, and they meet

the surface of another medium at any angle θ , then again we get reflection, but, in general, in a different direction. The law of reflection is that the reflected wave lies in the plane containing the incident wave and the normal to the surface at the point of incidence, but on the opposite side of the normal to the incident wave ; and its direction makes with the

FIG. 50

Reflection of Waves
 AB Reflecting Surface
 ON Normal to Surface
 PO Incident Wave
 OQ Reflected Wave
 $\angle PON = \angle NOQ$



normal an angle equal to that of the incident wave. This is shown in Fig. 50, where PO represents the incident wave, and OQ the reflected wave.

Interference

The formation of stationary waves by reflection is an example of a more general phenomenon known as *interference* of waves. If we have two wave-trains travelling through the same material in the same or different directions, and they happen to meet one another, then the displacement of any particle of the material will be that given by the resultant of the displacements which the two trains would give individually. Outside the region of interference each wave travels as though the other did not exist. The word "interference" is thus in one sense a misnomer. Neither wave interferes in the slightest degree with the other, but any particle in the path of the waves takes the movements imparted to it by both, and exhibits the resultant. Mathematically, the displacement of a particle is obtained by adding together the displace-

ments caused by the two wave-trains. We thus obtain, in the case of two "simple harmonic" wave-trains,

$$y = A_1 \sin 2\pi \left(n_1 t + \frac{x}{\lambda_1} \right) + A_2 \sin 2\pi \left(n_2 t + \frac{x}{\lambda_2} + a \right) \quad (13.12)$$

the term a being introduced to allow for the probability that the waves will not meet the particle at the same phase values. The displacement of a particle at any place and time can be obtained by inserting the appropriate values of x and t .

Forced Vibrations

A wave-train is the means by which energy is transmitted from a vibrating body to distant bodies. Let us now consider the effect on an initially stationary body, which is nevertheless free to move, of the successive impulses imparted to it by the wave-train—or, of course, if we wish, directly by the original vibrating body. We suppose that the receiving body, though free to move, tends when disturbed to return to its initial position.

In such a case the receiving body, when the first impulse reaches it, is displaced somewhat, and when the rarefaction succeeds the compression, it begins to return. At the next compression it is pushed away again, and again begins to return, and so on. Clearly, as a result of the treatment it receives, it also is set vibrating with the same frequency as that of the disturbing impulse. Such a vibration is called a *forced vibration*.

Now any body or mechanical system has a *natural frequency* or period of vibration, which is characteristic of itself. If, for instance, we clamp one end of a spring, momentarily displace the other end slightly, and then release it, the spring will quiver in a definite period, which is independent of how it is supported and remains the same as the vibrations gradually die away, *i.e.* as the amplitude of vibration diminishes. The same thing is true of a pendulum—in its simplest form a heavy bob suspended from a point by a fine thread. If l be

the length of the thread, the period of a complete to-and-fro vibration is $2\pi\sqrt{\frac{l}{g}}$, where g is the acceleration due to gravity. This period also is independent of the amplitude of swing (provided it is not very large) and of the magnitude of the weight.

Resonance

The result of applying to a system a periodic disturbance having the same period as that of its natural vibration is very interesting. Consider, for simplicity, the simple pendulum just referred to. Suppose a wave-train having the same period, T , is incident on the bob of the pendulum. The first compression displaces the bob slightly, after which it returns, swings past its equilibrium position, and then returns there again after a time T from the beginning. But at this moment the second impulse from the wave-train arrives, and it now meets a bob which has already a slight motion in its direction. It therefore causes a greater displacement than the first impulse. When the third impulse arrives, therefore, after another interval T , it strikes a bob which is moving still faster, and therefore sends it still further. Clearly, at each impulse the bob is given a greater amplitude than that which it had previously, and after a short time is set swinging violently. This phenomenon is called *resonance*. It occurs whenever a periodic disturbance is given to a system having the same natural vibration period.

There are many examples of resonance in everyday life. Thus, when a loud musical note is emitted in a room, a vase having the same natural period may be set vibrating, and in certain circumstances may even be broken if the disturbance is severe enough. Again, on walking over a suspension bridge, soldiers now break step, lest an accidental coincidence between the period of marching and that of swing of the bridge should set the latter vibrating so violently as to cause an accident. This actually occurred on one occasion.

EXERCISES

1. Describe the motion of the particles of a body through which (a) a longitudinal, (b) a transverse wave-train is passing.
2. Define the terms *velocity*, *wave-length*, and *frequency* in connection with a wave-train. Write down the equation of a simple harmonic wave-train, and explain the meaning of the various symbols which occur in it.
3. What is meant by "harmonic analysis"? Draw curves representing two simple harmonic wave-trains, of which the amplitude and frequency of one are respectively $\frac{1}{2}$ and $\frac{3}{2}$ those of the other, and find the resultant wave-train.
4. Explain how stationary waves are formed by interference between an incident and a reflected wave-train. What are the *nodes* and *loops* of a stationary wave?
5. Explain how forced vibrations are set up, and describe what happens when they are imparted to a body with a frequency equal to its natural frequency.

CHAPTER XIV

SOUND AND WAVES

Sound and Wave-motion

THAT the sensation of sound is caused by the impact on the ear of waves in the air is suggested by the fact that most sources of sound are found on examination to be vibrating bodies or systems of bodies. A violin string, for example, or the membrane of a drum, can easily be shown to be in vibration when giving out its note. The vibrations are imparted to the air surrounding it, which executes forced vibrations having the same frequency as the vibrating body, and are thus transmitted to the ear. In the common case of sound arising from instantaneous impact, such as that of a stone falling audibly on the ground, it is not so obvious that vibrations occur, but we may easily see that they do when we reflect that the molecules displaced at the point of impact return to their original positions and, as in all such restorations of displacement, overshoot the mark and oscillate to and fro for a short time. The fact that the oscillations are too rapid, and last for too short a time, to be easily observable does not alter the fact that they occur and can be detected with suitable instruments. Displacement of particles without vibration does not cause sound. An example of this is the silent motion of fish through water. Owing to the "streamline" form of the fish, the displaced water is able to find its way smoothly to the rear, and no to-and-fro oscillations are set up. Similarly, a slow compression and release of the molecules at the end of the rod

considered in the last chapter is inaudible ; it takes a sharp tap to make the rod vibrate.

Sound vibrations are of the longitudinal type. This must be so because they travel through gases. A gas cannot transmit transverse vibrations, for the molecules are so far apart that the sideways displacement of those in a given layer has little effect on those in the next layer, and energy therefore cannot be transmitted far in that way. Compression, however, is another matter. Gases are easily compressed, the molecules being moved together, and a compression is therefore handed on from layer to layer in the manner characteristic of longitudinal waves. We shall see also that the observed velocity of sound in air agrees well with the velocity calculated from the formula for longitudinal waves (p. 195).

Characteristics of Sound

Sounds may be distinguished by three characteristics—loudness (or intensity), pitch, and quality. These correspond respectively to the amplitude, frequency and form of the wave-train. A simple illustration of the dependence of loudness on amplitude is afforded by the fact that the harder we give our hammer taps to the rod, the louder is the note heard, though it remains in other respects (pitch and quality) the same. The harder blow corresponds to the greater displacement of the molecules, by Hooke's law, and therefore of the air set vibrating thereby.

It can be shown that pitch corresponds to frequency by measuring the frequencies of vibrating bodies by methods to be explained later, and comparing them with the notes heard. It is found that the higher the frequency the higher is the note. Since, by equation (13.1) a given frequency always corresponds to the same wave-length in a given medium (V , as we have explained, being constant for the medium), we may represent pitch by either frequency or wave-length. When a sound wave travels through a succession of different

media, however, it is the frequency and not the wave-length that remains the same throughout the journey. The association of a particular note with a definite frequency can be illustrated in a simple way by drawing a card across the teeth of a saw. When the motion is slow a succession of taps is heard, each corresponding to the transient vibrations of a single tooth as it is struck. With rapid movement, however, the taps may succeed one another so quickly as to give a note corresponding to the frequency of the taps, and this note rises as the speed of movement increases.

The quality of a sound (*e.g.* that which distinguishes a piano note from a violin note of the same pitch and intensity) can be shown to depend on the wave-form by means of special contrivances by which the wave-form is recorded. The simple sine form, represented by equation (13.3), is given by a properly struck tuning fork, but other musical instruments may give quite different forms. If a violin string is plucked or bowed at the middle, for instance, each half becomes a straight taut string, and this form is repeated at the extremities of each vibration. When the vibrations of a body are irregular, with no definite frequency or wave-length, we have a *noise*, as distinct from a musical note.

Whatever the form of a sound wave, it is usually possible to analyse it by Fourier's theorem into a number of simple harmonic vibrations, each of which could be produced by a tuning fork of a definite frequency. This analysis, however, is in general purely mathematical, and the component vibrations do not necessarily exist as objective sound waves. Sometimes, however, they do, and in that case the ear can detect them individually if they are intense enough. For example, in a gramophone record of an orchestral performance, the needle runs along a single groove whose form corresponds to the resultant vibration produced by all the instruments. Nevertheless, it is possible to pick out the notes of each instrument when the sound is reproduced.

Velocity of Sound

The fact that sound takes time to travel is a matter of common observation. Anyone who has seen a cricket match on a large ground, for instance, has noticed that he hears the sound of the bat striking the ball after he has seen it do so: this shows that sound travels more slowly than light.

It has already been said that the velocity with which longitudinal waves are transmitted through a medium is equal to the quantity $\sqrt{\frac{\text{Elasticity}}{\text{Density}}}$ for the medium. The elasticity, as we know, is the ratio of the force per unit area to the change of volume per unit volume. For a gas this is $-\frac{dP}{dV/V}$, where dP is the change of pressure in the medium corresponding to the compression or rarefaction in the wave (the minus sign is introduced because an increase of pressure causes a diminution—or negative increase—of volume, and *vice versa*). To evaluate this we must know the conditions under which the stress is applied, for the change of volume resulting from a given change of pressure depends on these. In a gas, for instance, we know that, for an isothermal change, $PV = \text{constant}$ and for an adiabatic change $PV^\gamma = \text{constant}$.

In the first case we have $PdV + VdP = 0$, i.e. $-\frac{dP}{dV/V} = P$.

For an adiabatic change, however, $-\frac{dP}{dV/V} = \gamma P$. It is the latter value that gives a velocity agreeing with the observed velocity of sound.

The reason for this can easily be understood. The successive compressions and rarefactions which constitute a sound wave-train take place very rapidly—some hundreds of times a second—so that there is no time for any heat developed in the process to be conducted away. The process is therefore an adiabatic one, and the temperature changes in accordance with the adiabatic equation. The measurement of the velocity

of sound in a gas, in fact, affords one of the best means of determining the value of γ for the gas (see p. 151).

For air at normal temperature and pressure (N.T.P.), $P = 1,013,250$ dynes per sq. cm., the density is 0.001293, and $\gamma = 1.4$. The resulting value for the velocity of longitudinal waves is 33,120 cm. (1,086 feet) per second, which is in good agreement with experimental values for the velocity of sound. This calculation has already been made (p. 151) for the determination of γ , taking the velocity of sound as known from direct experiment.

It will be noticed that the velocity is obtained independently of the frequency, or wave-length. This means that sounds of all frequencies travel with the same velocity. That this is so is a matter of common experience. When we hear a band playing at a distance, the sound reaches us some time after we see the conductor's beat, but all the notes played together arrive together, and there is no distortion of the music. The velocity of sound is independent of the intensity also ; this can be verified in the same way.

The principle on which determinations of the velocity of sound in open air are made is usually that of causing the sound and a light (or electric signal) to be emitted from a certain point at the same instant, and measuring the interval between their arrival at another point a known distance away. Light and electric signals travel at the rate of about 186,000 miles a second, so that their time of travel can be ignored. The measured interval can therefore be taken as the time of travel of the sound. It is on the same principle that, knowing the velocity of sound, we can calculate the distance of a lightning discharge by noting the interval between the flash and the thunder.

In such experiments the effect of wind is important. The measured velocity is the resultant of the velocity of sound and whatever wind may be moving the air through which the sound passes. Thus, if the sound is travelling with the wind, and the measured velocity is V while the velocity of the

wind is v , the velocity of sound in still air is $V - v$. If the sound is travelling against the wind, the velocity in still air is, of course, $V + v$.

The formula already given for the velocity of waves in a medium shows that that velocity, when the medium is a gas, should be independent of the pressure. For, at constant temperature, $\frac{P}{\rho}$ is constant, by Boyle's law (p. 120), where

ρ is the density, and the velocity of sound is $\sqrt{\frac{\gamma P}{\rho}}$. The value of $\frac{P}{\rho}$ is, of course, $R\theta$ (p. 126), where θ is the absolute temperature, so the velocity of sound should be given by $\sqrt{\gamma R\theta}$, and should therefore vary with temperature. This agrees with observation. This result, of course, does not apply if the composition of the atmosphere changes; for example, by variations in humidity.

The velocity of sound in other than gaseous media is given in terms of the elasticity and density by the same general formula. With liquids, for instance, the density is much greater than with gases, but the elasticity is magnified still more, so that sound travels considerably faster in liquids than in gases. The value for water, measured at about 8° C. , is found to be 143,500 cm. a second, agreeing with the calculated value to about 4 per cent. This value is more than four times the velocity of sound in air.

One point is worth noting here. When a sound wave passes from one medium to another—say from air to water—the velocity changes. Hence, from equation (13.1), the product $n\lambda$ changes. Now the frequency must be the same throughout the wave-train, or the energy would quickly be dissipated: we may verify this by the fact that the same note is heard, whatever the medium. Hence the wave-length must change. Clearly, it is much greater in water than in air.

When a gas is enclosed in a tube, the velocity of sound in it is not necessarily the same as in the free medium. The

appropriate modulus of elasticity is clearly something between the bulk modulus and the analogue for gases of Young's modulus, so we should expect a different value for the velocity. There are, however, other complications also. It appears that with air in tubes the velocity decreases slightly with decreasing intensity of the sound, and increases with increase of the diameter of the tube. The variation is not great, its extreme range being of the order of 1 per cent., but it is quite definite.

Location of Sound

An interesting application in warfare of our knowledge of the velocity of sound is found in a method used for the location of enemy guns. Suppose a gun is situated at G (Fig. 51),

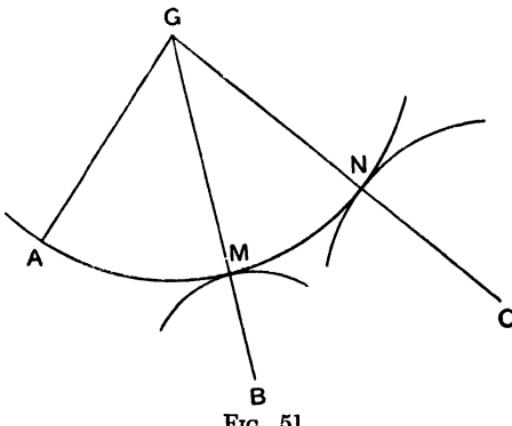


FIG. 51
Location of Guns by Sound

G Gun

A, B, C Observing Stations

and observers at A, B, and C note the times at which it is heard to fire. With centre B and radius equal (on any convenient scale) to the interval between the times of hearing the shot at A and B, draw a circular arc ; and with centre C and radius equal to the interval between the times of

hearing the shot at A and C, draw a second arc. Now draw a circle tangential to these arcs, and passing through A. The centre of this circle will be the position of the gun.

This is at once clear from the diagram. For if G be the origin of the sound, the distances GA, GB, and GC are respectively proportional to the times of travel of the sound to stations A, B, and C. GA, GM, and GN are equal, being radii of the same circle, so that MB and NC must be equal to the lag in the arrival of the sound at B and C, respectively, with respect to A. But this is so, since we have drawn the arcs to satisfy this condition. Furthermore, there is no other point from which lines can be drawn to A, B, and C to satisfy this condition. Hence the gun must be at G.

EXERCISES

1. Give reasons for believing that sound is caused by the impact on the ear of waves in air caused by vibrating bodies, and state what physical characteristics of the sound correspond to the amplitude, frequency, and wave-form of the air waves.
2. An aeroplane travelling horizontally at 240 miles per hour drops a bomb. Neglecting the resistance of the air, calculate the height of the aeroplane if the bomb and the sound of its release reach the same point on the ground at the same moment. Find also the time of fall. The velocity of sound may be taken as 1,100 ft. per second.
3. At 4° C. the adiabatic elasticity of water is 2.03×10^{10} and its density is 1. Find how long a sound wave will take to cover a distance of one kilometre in water at this temperature against a current of 30 cm. a second.
4. Explain how the ratio of the principal specific heats of a gas may be determined from measurements of the velocity of sound in the gas.

CHAPTER XV

THE FREQUENCY OF SOUND WAVES

Measurement of Frequency

A SIMPLE method of measuring the frequency of a tuning fork is to attach a light style to the fork and let it rest on a long flat plate moving uniformly in a direction perpendicular to that of the vibrations. If the plate is smoked, so that the style removes a line of soot where it makes contact with the plate, we see the wave form of the vibration traced out on the plate. If the speed of the plate is known, we have then simply to count the number of waves in the distance travelled in one second, in order to obtain the frequency. A convenient way of moving the plate is to let it fall under gravity, but in that case, of course, its velocity is not constant, and we must allow for the acceleration.

Another very interesting way of measuring frequency is by the "stroboscopic" method. The principle of this method is that if a uniformly moving row of similar objects, such as a row of dots on a card, is viewed intermittently and appears to be at rest, the interval between two successive views of it must be an integral multiple of the interval between the passage of adjacent dots across the line of sight. An example of this is sometimes seen at cinematograph performances when a moving wheel appears on the screen. If the interval between the successive stationary pictures of which the apparently continuous scene is composed is an integral multiple of the time taken by one spoke to move into the position occupied by its neighbour, the wheel appears to be at rest, for one spoke is indistinguishable in appearance from

another, and the successive appearances of the wheel seem to be identical.

To apply this to the measurement of frequency of a tuning fork, say, two light cards are fixed to the prongs,* having slots in them which, when the fork is in its equilibrium position, are directly opposite one another (Fig. 52). An observer can then see through the slots, but cannot do so when the prongs are in any other position (*i.e.* at any other phase of the vibration). Behind the cards is placed a disc, having marked round its circumference a succession of equally

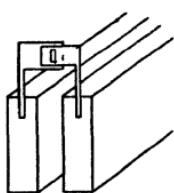


FIG. 52

Tuning fork arranged for stroboscopic observation

spaced dots, in such a position that when the disc is rotated about its centre the dots pass successively behind the slots through which the observer is looking. Now suppose a particular dot is seen through the slots when the fork and disc are both at rest. On setting the fork vibrating, this dot will be seen intermittently, but as the vibrations are rapid it appears to be seen continuously, and, of course, since the disc is at rest it appears to be, as it is, stationary. If, now, the disc is made to rotate slowly, the motion will be visible because, during the time of one vibration of the fork, the disc has rotated so that the space between the dots is seen, and several vibrations will have to occur before the next dot comes into the line of sight. On increasing the speed of the disc, however, a time will come when the interval between two successive views through the slots is just equal to the time required for one dot to move into the position of its neighbour, and then the disc will appear stationary again. If we know the speed of rotation of the disc, we can then determine the frequency of the fork in the following way.

The interval between two views through the slots is half

* Whenever anything is attached to the prong of a tuning fork it must be as light as possible, otherwise the frequency of vibration will be appreciably altered by its presence.

the frequency of the fork $\left(\frac{n}{2}, \text{ say}\right)$, because a complete vibration consists of a to-and-fro motion of the prongs. Suppose there are m dots equally spaced along the circumference of the disc, which we will suppose makes R revolutions a second —*i.e.* one revolution in $\frac{1}{R}$ second. Then the time each dot takes to move into the position of its neighbour is $\frac{1}{Rm}$ seconds. Equating this to $\frac{n}{2}$ we can determine n .

Experiments such as these show that the frequency of the middle C (written c') in the ordinary diatonic scale is 256 vibrations a second. Octaves of this note correspond to integral multiples of this frequency; thus, the next higher C note (c'') has 512 vibrations a second, and so on. The various notes of the scale have frequencies which bear simple ratios to these numbers. Thus, within the octave c' to c'' , the ratios of the vibration numbers for the successive notes (the white notes on the piano) are 24, 27, 30, 32, 36, 40, 45, 48; and the same ratios hold for any octave extending from a C to a C note.

Concord and Discord

This gives a clue to the physical counterpart of concord and discord in music, *i.e.* to our sensations of pleasure or displeasure when two notes are sounded together; it is roughly that the concord is greatest when the ratio of the frequencies of the notes is simplest. Thus, if we use tonic sol-fa notation, and represent the scale just given by d, r, m, f, s, l, t, d' , we see that the frequency ratios of d to m, f, s, l , and d' are respectively $\frac{4}{3}, \frac{3}{4}, \frac{2}{3}, \frac{3}{5}$, and $\frac{1}{2}$; and the combination of any of these notes with d is pleasing. The ratios of d to r and t , however, are the less simple fractions $\frac{8}{9}$ and $\frac{8}{7}$, and the combinations are less pleasing.

A more exact statement can be given in terms of *beats*, which occur when two discordant notes are sounded together.

These arise in the following way. The resultant of two vibrations may be obtained by adding the ordinates of the two wave-forms, as in equation (13.12). For simplicity, let us consider a less general case, in which we put $x = 0$ (*i.e.* we consider only a particular point in the wave-trains), and take the amplitudes and phases of the waves to be the same. The resultant vibration is then given by

$$\left. \begin{aligned} y &= A \sin 2\pi n_1 t + A \sin 2\pi n_2 t \\ &= 2A \cos 2\pi \left(\frac{n_1 - n_2}{2} \right) t \cdot \sin 2\pi \left(\frac{n_1 + n_2}{2} \right) t \end{aligned} \right\} \quad (15.1)$$

If, now, we write

$$\left. \begin{aligned} A' &= 2A \cos 2\pi \left(\frac{n_1 - n_2}{2} \right) t \\ \text{and } p &= \frac{n_1 + n_2}{2} \end{aligned} \right\} \quad . . . \quad (15.2)$$

the resultant becomes

$$y = A' \sin 2\pi pt, \quad . . . \quad (15.3)$$

and this may be regarded as a single vibration with frequency p and amplitude A' .

Now the amplitude itself varies periodically, as can easily be seen, for from (15.2) it vanishes when

$$2\pi \left(\frac{n_1 - n_2}{2} \right) t = \left(q + \frac{1}{2} \right) \pi \quad . . . \quad (15.4)$$

where q is an integer; *i.e.* when

$$t = \frac{q + \frac{1}{2}}{n_1 - n_2} \quad . . . \quad (15.5)$$

Giving q two successive integral values, we have then

$$\left. \begin{aligned} t_1 &= \frac{q + \frac{1}{2}}{n_1 - n_2} \\ t_2 &= \frac{q + \frac{3}{2}}{n_1 - n_2} \end{aligned} \right\} \quad . . . \quad (15.6)$$

so that

$$t_2 - t_1 = \frac{1}{n_1 - n_2} \quad . . . \quad (15.7)$$

This is the interval between two successive zero values of the amplitude, and is therefore the period of variation of the amplitude. The corresponding frequency of the amplitude is, of course, $n_1 - n_2$.

Now suppose n_1 and n_2 are not very different ; then this frequency is small, so that instead of observing it as a "musical" variation, so to speak, we hear the individual pulses, or "beats," separately, just as, in our previous illustration (p. 211) of the card drawn along the teeth of a saw, the separate impacts could be heard for small frequencies and the musical note only for large ones. Remembering that the amplitude measures the intensity of the sound, we see that this periodic vanishing means that we hear the sound sinking periodically almost to nothing, with intermediate rises to maximum strength, the frequency of the note heard, of course, being p .

Now, a discordant combination of notes corresponds to one in which the frequency of the amplitude variation, $n_1 - n_2$, is small. In Fig. 53 are drawn the resultant curves for the combinations d and s (frequency ratio $\frac{2}{3}$) and d and r (frequency ratio $\frac{8}{9}$). A variation of amplitude is shown in both cases, but in the latter it is more pronounced, though its frequency is the same as in the former ; the beats are then sufficiently near distinct audibility to cause unpleasantness.

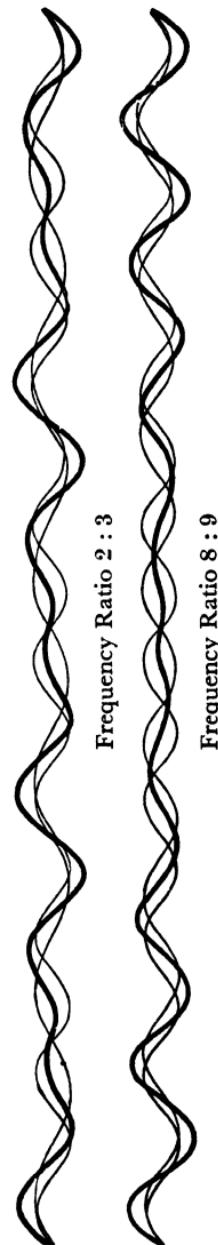


Fig. 53.—Resultants (heavy curves) of superimposed waves having different frequencies

Overtones and Harmonics

We have supposed here that only two vibrations are compounded, and if we have tuning forks sounding them, this corresponds to what actually occurs. But with most musical instruments we do not get only single tones having one definite frequency; accompanying them are a series of *overtones*, as they are called, which are vibrations of higher frequency than the chief, or *fundamental* one. It often happens that the frequencies of the overtones are integral multiples of the fundamental frequency, in which case they are called *harmonics*. When this occurs, clearly the components into which the complete sound emitted can be analysed mathematically by Fourier's theorem correspond to actually existing sound waves. Thus, when the piano note c' is struck, a note with the corresponding frequency is produced as the fundamental tone, but accompanying it there are notes c'' , g'' , c''' , etc. When two notes are sounded together, therefore, we have to take into account the possibility of beats not only between the fundamentals but also between the overtones. The most concordant notes are those which have overtones in common, so that for these the possibility of beats is excluded.

Range of Audibility

The range of frequencies which the ear can detect as sound is limited at both ends. As we have seen, a succession of taps whose frequency is regularly increased passes gradually into a musical note, and it is impossible to state precisely the point at which the change takes place. The matter is complicated by the fact that overtones, which have higher frequencies than the fundamental note, may be heard and mistaken for the fundamental. Careful experiments, however, have shown that the lower limit of frequency for a definite note to be heard is about 30 a second, but the number varies for different persons. These low frequencies corre-

spond to very low notes of the scale, the note rising with the frequency.

At the other end of the scale it is found again that there is considerable variation from one person to another. We may take 24,000 vibrations a second as not far from the upper limit for human beings, though many find that their limit lies considerably below this. Some animals, on the other hand, can probably hear higher notes.

Supersonics

Vibrations of higher frequency than those which are audible are known as *supersonics*; in recent years they have received numerous applications. The effective production of supersonics depends on the fact that certain crystals—particularly quartz—change in length when placed in an electric field. There is a kind of invisible “grain” in quartz, the direction of which is known as the *optic axis*, and if a crystal be cut parallel to this axis, and electrodes from a battery be connected so as to apply a field perpendicular to it (Fig. 54), the length of the crystal is changed in the direction perpendicular to both the optic axis and the field.* If, now, the direction of the field be made to alternate rapidly, the crystal will lengthen and shorten rapidly also; and if the period of the alternations be made equal to the natural period of vibration of the crystal, the amplitude of the oscillations will be considerable, and waves will be sent out into the surrounding air.

The velocity of sound in quartz is 5,500 metres a second, and it is easy to see that the wave-length of the vibrations is twice the thickness of the crystal, for the directions of motion

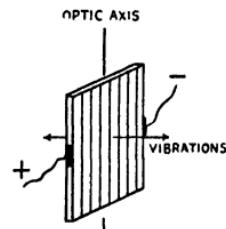


FIG. 54

Direction of vibration of quartz crystal in alternating electric field

* The volume of the crystal does not change. The dimension in the direction of the field varies also, in such a way that the volume remains constant.

at the two surfaces as the crystal expands and contracts are in opposite directions. Consequently, for a frequency of 55,000, say—giving, from equation (13.1), a wave-length of 10 cm.—we need a crystal 5 cm. thick. Thinner crystals clearly give higher frequencies.

Such oscillations have been used to control clocks ; they give very accurate measures of short intervals of time. They are effective also in breaking up colloidal substances into fine drops, and in changing the forms of certain chemical molecules ; for example, in reducing starch to a form of sugar. Several other applications of supersonics have been made, and the range of their usefulness is likely to widen in the near future. The energy in these vibrations is considerable, and small creatures, such as certain fish, have been killed when subjected to them.

Vibrating Strings

Many interesting phenomena are exhibited by the vibrations of strings. To study these it is convenient to use the *monochord* ; this consists of a wire fixed at one end, A, (Fig. 55) which

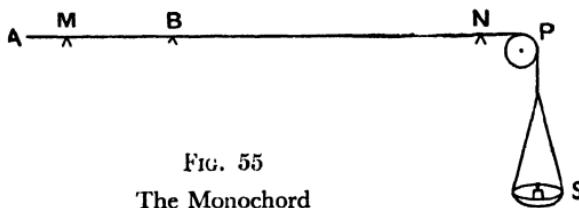


FIG. 55
The Monochord

passes over two fixed bridges, M, N, and a pulley, P, to a scale-pan, S, on which weights can be placed. A movable bridge, B, can be placed anywhere along the wire so as to vary the length (namely, that between two consecutive bridges) which is set in vibration.

Now it can be shown that if such a wire is plucked and then released, the frequency, n , of the transverse vibrations which it makes depends in a perfectly definite way on the stretching force (T), the mass (m) of unit length of the wire,

and the total length (l) which vibrates. For it is found that the velocity with which such a vibration would travel along the wire if it were free to do so is given by

$$V = \sqrt{\frac{T}{m}} \dots \dots \dots \quad (15.8)$$

an obvious analogue of the expression for the velocity of longitudinal waves in an unbounded medium. But since the ends of the wire are fixed, and are therefore always nodes in the vibration, the distance, l , between them must be half a wave-length, *i.e.* $\frac{\lambda}{2}$. Hence, from our fundamental equation (13.1) we have

$$n = \frac{1}{2l} \sqrt{\frac{T}{m}} \dots \dots \dots \quad (15.9)$$

This can be verified with the monochord, for T depends on the weights in the pan, m and l can be measured and varied as we wish, and the frequency can be determined by comparing the note emitted with the notes of a set of tuning forks of known frequency. It is preferable to have a sounding box under the stretched string in order to amplify the sound produced.

A violin is, in principle, a set of four monochords, in which, however, only slight variations of length are possible by changing the position of the bridge. Tuning is carried out by varying T —the ordinary process of tightening or loosening the string. This also gives only slight adjustments of frequency, and the chief difference between the four strings is in the value of m , the mass per unit length.

Melde's Experiment

The frequency, n , is that of the fundamental note given by the string, but there are also overtones, which in this case are harmonics. This may be shown most strikingly by what is known as *Melde's experiment*, illustrated in Fig. 56. In this

experiment the string is attached at one end to a prong of a tuning fork, arranged so that the vibrations are in the direction of, or perpendicular to, the length of the string. If the fork is made to vibrate,* then, the weight being W , the string will vibrate with the frequency of the fork. Now suppose the weight is reduced to a quarter of its value. Then, by formula (15.9), the same frequency corresponds to a string of half the length. It is then found that the string vibrates in two parts, as shown in the figure, *i.e.* it has a node in the middle, as well as at the ends, and this point remains at rest throughout the

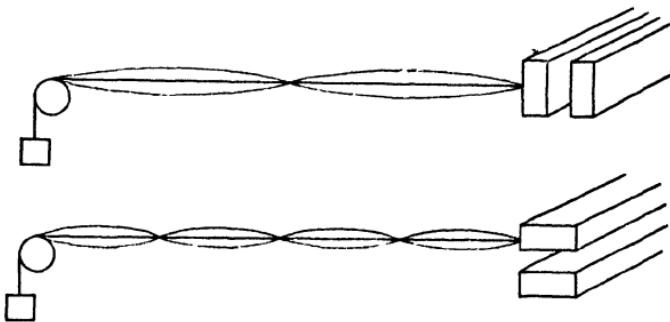


FIG. 56
Melde's Experiment

vibrations. Similarly, we can make $\frac{1}{3}$, $\frac{1}{4}$. . . of the string vibrate by suitable adjustment of the weight. These vibrations constitute the harmonics of the fundamental one.

There is a difference between the results in the two cases illustrated in the figure. In the first case we can see that when the fork is at its outer extremity, the string is loose ; and half a vibration later, when the fork is at its inner extremity, the string is taut. This change in the string corresponds to a quarter of a complete vibration, so that the period of the string is twice that of the fork. In the lower figure, however, the string is taut at both extremities of the vibration of the fork, so that here the periods of string and fork agree.

* It is a common practice to maintain a tuning fork electrically, supplying energy so that the vibrations are not damped out (see p. 237).

It is not necessary, however, to establish the special conditions of Melde's experiment in order to obtain the harmonics of a string. They are present, together with the fundamental note, even with the ordinary monochord. This may be shown by setting the string in vibration, and then lightly damping it at a point which is a node for a harmonic but not for the fundamental, *e.g.* at a quarter of the way along the string. Such damping quenches the fundamental, but has no effect on the harmonic, because the point of damping, being a node, is at rest in any case. The result is that the fundamental note dries out, and the harmonic, which was previously drowned, is heard.

Resonators

This is made much more evident if a *resonator* is used to amplify the sound of the harmonic. A resonator is some mechanical system which has a natural period of vibration equal to that which it is desired to detect. The existence of a vibration with that period then makes the system vibrate by resonance, and if it is of such a form that the sound can be conveniently heard, its purpose is fulfilled. A suitable form of resonator, which was devised by Helmholtz, is a hollow globe, usually made of thin metal, with a wide opening on one side, and on the other a narrow opening which can be placed in the ear (Fig. 57). The natural period of the globe will depend on its dimensions and on the size of the wide opening, and a series of such resonators can be kept, which resound to different frequencies. When a particular faint note is to be heard, the ear is placed against the corresponding resonator.

A simple example of a resonator—which, however, is not so effective as that just described—is an open cylinder,

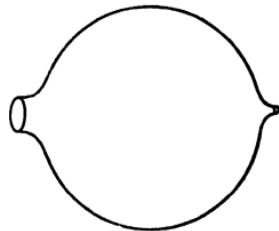


FIG. 57

Helmholtz Resonator

the effective length of which can be varied by pouring in water. The natural period varies with the depth of air, as may easily be noticed by the variation in the sound made by the pouring as the water rises in the cylinder. When a tuning fork is set vibrating and held over the resonator having its own frequency, the sound swells out considerably.

The Doppler Effect

We have so far supposed that the frequency of a sound is the same as that of the vibrating body which originates it. Strictly speaking, this is not necessarily true : the frequency of the sound which is heard is that of the vibrations falling on the ear, and this is the same as that of the vibrating source only if the source and hearer are not approaching or receding from one another. If they are approaching, the note heard is higher than that emitted, and if they are drawing apart, it is lower. This is known as the *Doppler effect*. It is very noticeable when a train passes an observer while blowing its whistle ; an abrupt fall of pitch is heard as the whistle ceases to approach and starts to recede.

The amount by which the pitch is changed depends on the frequency of the note, and also on the velocity with which the distance between source and observer is changing. It is entirely a relative matter, and has nothing to do with the motion of either source or observer alone. Thus, an observer on the train which is blowing the whistle, though both observer and whistle are moving, hears the same note as he would if both were at rest, because the effect depends only on the relative motion between them. Again, if the observer approaches or recedes from the source with velocity v , the effect is the same as if the source approaches or recedes from the observer with velocity v . We can easily derive an expression for the change of pitch in each case, but for want of space we will take one case only—that in which an observer is approaching a stationary source with velocity v .

In Fig. 58 let S be the source emitting sound waves which

travel with velocity V cm. per second, and let O_1 and O_2 be two observers who, at a certain instant, t , are at the same point, O_1 , distant V cm. from S . Let O_1 be stationary with respect to the source, while O_2 is moving towards the source with velocity v . At time $t + 1$, *i.e.* one second later than the instant at which O_1 and O_2 were together, O_2 will be at the

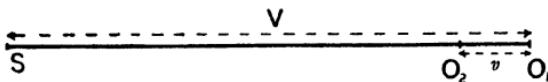


FIG. 58

The Doppler Effect

point shown on the figure, distant v from O_1 . Now if n be the frequency of the waves emitted from S , O_1 will receive n waves per second, and will therefore hear a note of frequency n . O_2 , however, during the same second will receive all the vibrations which arrive at O_1 , and, in addition, the vibrations contained in the distance $O_1 O_2 (= v)$, which have not yet reached O_1 . The wave-length of the emitted vibrations being $\frac{V}{n}$, the number in the distance v will be $\frac{v}{V/n} = n \frac{v}{V}$. Hence the vibrations received by O_2 in one second will be $n + n \frac{v}{V} = n \left(1 + \frac{v}{V}\right)$. O_2 will therefore hear a note of this frequency, which is clearly higher than n .

The same formula applies to all cases, but when the source and observer are receding from one another we must change the sign of v , and we obtain $n \left(1 - \frac{v}{V}\right)$ for the frequency heard.

This effect has no important applications in sound, but the corresponding effect with light is of the greatest importance, for it enables us to determine the speed with which we are approaching or receding from various stars—a quantity of which we should otherwise be in complete ignorance.

EXERCISES

1. Describe the stroboscopic method of measuring the frequency of a tuning fork. What would you expect to observe if the rotational speed of the disc used in the experiments were doubled ?
2. Draw the resultant of two simple harmonic vibrations of equal amplitude, of which the frequency of one is twice that of the other.
3. Explain the occurrence of beats when two notes of slightly different pitch are sounded together.
4. What frequency of supersonic waves would you expect from the resonant oscillations of a quartz crystal 1 cm. thick ? Explain how such oscillations may be maintained, and state some of their applications.
5. A horizontal steel wire is stretched by a weight of 6 kilograms hanging over a frictionless pulley, and a length of 1 metre between two fixed points is bowed. What is the frequency of the fundamental note produced if the diameter of the wire is 0.2 mm. and its density 7.7 ?
6. Describe and explain the Doppler effect. Find the speed with which two trains must pass one another in order that the note of a whistle blown by one shall appear to an observer on the other to change on passing from c'' (frequency 512) to g' (frequency 384).

CHAPTER XVI

THE REFLECTION OF SOUND

Examples of Reflection

SOUND waves, like waves in general, can be reflected from the surface of separation between two different media according to the law already stated (p. 205). This can be verified very simply by placing two tubes, as in Fig. 59, equally inclined on opposite sides of the normal to a solid reflecting surface, and placing a source of sound at S and the ear or a resonator at E. Direct passage of the sound from S to E can be minimized by interposing some absorbing material. It is then found that the sound heard is much more intense than when the tubes make slightly different angles with the normal.

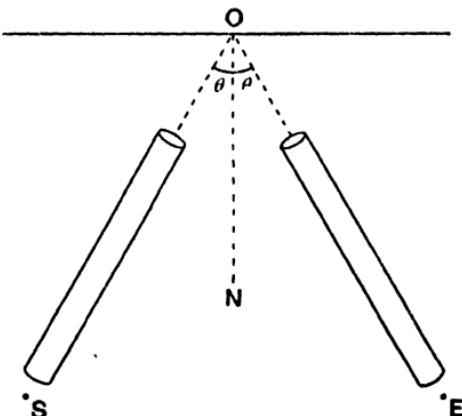


FIG. 59
Reflection of Sound

A natural example of the reflection of sound is provided by echoes. The note reflected from a distant hillside is heard some time after the original note, owing to the fact that sound takes time to travel to the hillside and back. Whispering galleries, such as that in the dome of St. Paul's Cathedral, afford

another example. The sound issuing from a point just inside the circumference of the circle is reflected successively along a series of chords, and is thus prevented from spreading out into space. Reflection from the walls of buildings is apt to cause difficulties in hearing unless precautions are taken against it. The poor audibility in many halls is a consequence of the confusion between the sound which comes direct to the hearer and that which is reflected and reaches him slightly later. In ear trumpets the reflection of sound is turned to advantage by enabling the ear to receive not only the sound which would normally enter it, but also that reflected from the sides of the trumpet.

Sound Waves in Pipes

When the sound is confined within a pipe, reflection from the far end plays an important part in the behaviour of the pipe as a resonator. Suppose we have a pipe of length l (Fig. 60), over which we place a vibrating tuning fork. To some extent the pipe will augment the sound whatever the relation between its length and the frequency of the fork, owing to the forced vibrations set up in the air filling it, but for each frequency there is a particular length for which the pipe resounds most strongly, and the augmentation is then very great.



FIG. 60
Resonance
in Pipes

We can easily calculate what this length must be. Suppose the lower prong of the vibrating fork in the figure is at the upper extremity of its vibration, and is just about to descend. On doing so it compresses the air in front of it, and this compression travels down the pipe and, supposing the bottom of the pipe to be a rigid surface, is reflected back again as a compression. If, now, it arrives back just when the prong of the fork has reached its lowest position, it will assist the return journey of the prong, so that the vibrations of the air in the pipe will synchronize with those of the fork, and we

shall have resonance. The condition for this is that the time taken for the sound to travel the distance l and back (*i.e.* a total distance $2l$) shall be that taken by the fork to execute half a vibration. If the velocity of sound in air is V , the time taken by the air wave is $\frac{2l}{V}$, and if the frequency of the fork is n , the time of half a vibration is $\frac{1}{2n}$. Hence, equating these two quantities we have

$$l = \frac{V}{4n} \dots \dots \dots \quad (16 \cdot 1)$$

Now $V = n\lambda$, where λ is the wave-length of the vibration in air. Hence

$$l = \frac{\lambda}{4} \dots \dots \dots \quad (16 \cdot 2)$$

i.e. the length of the pipe is one-quarter of the wave-length of the note to which it will resound.

This result is not quite in accordance with observation, because the length for resonance depends partly on the diameter of the pipe. It can be understood that the motion of the air in the pipe will be determined to some extent by the change from the confining walls to the freedom of the open air (we have already seen (p. 214) that the velocity of sound in open air is not the same as that in pipes), but the exact magnitude of this effect is difficult to calculate. Experiments show, however, that instead of l in (16·2) we must take approximately $l + 0 \cdot 6r$, where r is the radius of the pipe. We have then

$$\lambda = 4(l + 0 \cdot 6r) \dots \dots \quad (16 \cdot 3)$$

as the wave-length in air of the note to which the pipe resounds. This adjustment of the length is generally known as the *end correction*. Having pointed it out, we shall henceforth ignore it, for simplicity of presentation, but it must be remembered that it should always be taken into account in practical applications.

Stationary Waves in Pipes

In deriving the formula for λ we have considered only one-half of a vibration of the fork. In the succeeding half, essentially the same thing is repeated, but now we have a rarefaction reflected at the bottom of the pipe, and in subsequent vibrations the same sequence of events is repeated over and over again. We have here an example in the pipe of stationary waves formed by the interference of the reflected with the incident wave. At the bottom of the pipe, which is rigid, the air is stationary, so that this point is a node; and at the top, where the motion is a maximum, we have a loop. The vibration in the pipe is represented, in fact, by Fig. 48, in which the wall represents the closed end and AB the open end. The length of the pipe is clearly a quarter of a wave-length.

We can get resonance not only from such a pipe as that already considered, but also from one which is open at both ends, but in this case the wave-length for resonance is $\frac{\lambda}{2}$

instead of $\frac{\lambda}{4}$. The reason for this is the sudden change of phase on reflection at a yielding surface (see p. 204). This causes a loop to form at each end of the pipe, which is now represented by the distance between AB and the wall in Fig. 49.

In both cases the fundamental vibrations so far discussed are accompanied by overtones. If we remember the necessary conditions that with a pipe closed at one end the closed end must be a node and the open end a loop, and that with a pipe open at both ends each end must be a loop, we can see that in the first case only the odd harmonics, and in the second case only the even harmonics will be present. For (see Fig. 48) the next possible length for which a pipe closed at one end will resound to the same note as that already considered, is equal to the distance from CD to the wall, which is three times the original length, and if the fundamental

note of such a pipe be λ_1 , say, then the original wave-length (now the first harmonic) is $\frac{\lambda_1}{3}$. Similarly we can show that

$\frac{\lambda_1}{5}, \frac{\lambda_1}{7}, \dots$ are possible resonance notes. A similar argument, in terms of Fig. 49 instead of Fig. 48, shows that the resonance notes for a pipe open at both ends are λ_2 , $\frac{\lambda_2}{2}, \frac{\lambda_2}{4}, \dots$. All this can be verified by experiment.

The most important practical application of this phenomenon is, of course, to organ pipes, which may be of either of the two kinds considered. The exact conditions at the ends have, of course, to be carefully considered here, and the foregoing discussion concerns only the fundamental principles.

Kundt's Dust Tube

The vibrations in a pipe closed at one end can be exhibited very beautifully by a method devised by Kundt, and illustrated in Fig. 61. A glass tube has at one end, A, a tightly

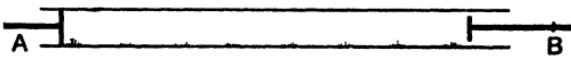


FIG. 61

Kundt's Dust Tube

fitting piston, which may be firmly fixed at any desired position along the tube. At the other end is a similar piston, not touching the walls of the tube, attached to a rod of glass or metal which is fixed at some point B. This rod can be set in longitudinal vibration by stroking it with a cloth. For glass tubes a wet cloth is most suitable, while for metal ones resined leather gives the best results. The vibrations are communicated to the air in the tube, and A is moved until a position of resonance is found—indicated by the sudden swelling out of the sound—when it is tightly clamped. This

position will normally correspond to an overtone for the pipe, with apparatus of ordinary laboratory dimensions. A is a closed end, while B is intermediate between a closed and an open end. We shall thus have a node at A and another not far from B, since B is neither a node nor a loop. Other nodes and loops will occur at intermediate points.

The essential feature of Kundt's tube is the method of showing these facts visually. A very fine and light powder—lycopodium is generally used—is placed in the tube, and when the vibrations occur the motion of the air causes it to rise up as though blown about inside the tube. When the vibrations cease, it settles again, but mainly at points where there is no motion of air, *i.e.* at the nodes. We thus observe it collected in little heaps, the distance between neighbouring heaps being half the wave-length of the vibrations.

This affords a very convenient method of measuring the velocity of sound in pipes for different gases. The frequency, n , of the vibrations is, of course, that of the piston fixed at B, and this can be determined by matching the note heard with that of a suitably adjusted monochord or tuning fork. The wave-length, λ , in the gas is twice the distance between adjacent heaps of powder, and a fairly accurate value can be found by taking the average of such distances. We thus obtain the velocity, V , from the relation $V = n\lambda$.

EXERCISES

1. Describe the formation of stationary sound waves in pipes, and calculate the length of a pipe closed at one end which will resound to a tuning fork of frequency 256, if the end effect can be neglected.
2. Explain the principle of Kundt's dust tube. How would you use it to determine the velocity of sound in a gas?

CHAPTER XVII

PRODUCTION AND DETECTION OF SOUND

Sources of Sound

SOURCES of sound are so common and various that it scarcely seems necessary to give special consideration to the production of sound. In experimental work, however, it is usually important that the sounds produced shall remain uniform in each characteristic, and it is often desirable that they shall be pure fundamental tones, without overtones. We can only touch on the subject here, but it is clear that any device adopted should be as simple as possible, so that it can be kept under perfect control. A tuning fork, if not struck violently, gives a very fine tone, and it is usually kept vibrating by an electrical circuit completed and broken periodically, as in the ordinary electric bell. The current activates an electro-magnet, which attracts a prong of the fork each time the current flows and releases it when the current stops. By arranging that the frequency of breaking of the current shall be equal to the natural frequency of the fork, a strong, permanent, pure note is obtained.

A source of sound now only too familiar is the *siren*. This consists of a rotating disc, pierced round the circumference by a number of equidistant holes, behind which a fixed nozzle emits a constant jet of air. As each hole comes in front of the nozzle a puff of air goes through, and this, repeated many times a second, gives the alternate compressions and rarefactions which form a sound wave. The quality of the sound depends on the form of the nozzle and holes. If there are n holes, and the disc rotates N times a second, the frequency

of the note is nN . A resonator tuned to this frequency can be placed on the opposite side of the disc to the nozzle, in order to swell the sound. By varying the speed of rotation (*i.e.* N), the well-known wailing sound is produced.

Many other sources of sound are employed in modern work, most of which involve electrical principles. The vibrating crystals already mentioned are very satisfactory, but are used mainly for supersonic vibrations.

Amplification of Sound

The amplification of sounds is as important as their production. The principle of resonance is employed, a simple example—Helmholtz's resonator—having already been mentioned. More modern instruments, such as the thermionic valve and the microphone, will be described in the second volume of this work, when the principles of electricity have been discussed.

Detection of Sound

The simplest and most convenient instrument for the detection and differentiation of sound is the ear. This organ is a much more satisfactory scientific instrument for the purpose than any artificial means so far devised. This is not the case with all the senses. For instance, the detection of temperature difference by direct feeling has been almost entirely superseded by the use of the thermometer, while the photographic plate has rather less completely but still largely displaced the eye as a detector of light ; but for sound we still depend mainly on the ear. This is not only a matter of sensitiveness. The ear also has analytical properties not possessed by the eye, for example. When the eye is presented with a mixture of colours, it sees merely a single colour—the resultant. It is not able to say what colours have gone to the production of this resultant. When the ear is presented with a mixture of sounds, however, the individual sounds can be distinguished and their pitch, intensity, and quality can

all be observed independently of the corresponding characteristics of the other components of the mixture.

Sometimes, however, it is inconvenient or impossible to apply the ear to the detection of sound, and other means have to be used. One of the simplest of these is the *manometric flame*, introduced by Koenig, and used mainly, though by no means entirely, to distinguish the variations of pressure of the air vibrating in an organ pipe. This device consists of an ordinary gas burner, in which, however, the gas on its way to the jet passes through a small chamber closed on one side by an indiarubber membrane, the outer surface of which can be exposed to the source of sound. The membrane responds to the successive compressions and rarefactions forming the sound wave, and the flame accordingly jumps up and down. Such a flame can be a very sensitive indicator of the existence of sound waves, responding especially well to the sibilants when one speaks near it. The strength of its response gives some idea also of the intensity of the sound. By placing such burners at different points along an organ pipe and examining the fluctuations of the flame by a revolving mirror so that the rapidly changing forms can be seen separately, we get a very useful exhibition of the variation of pressure inside the pipe.

Location of Submarines : An interesting application of the detection of sound was made during the war of 1914-18 to the location of submarines. The sound waves emanating from a submarine travel through the water and afford a means of locating the vessel if they can be detected sufficiently well. The ear is unsuitable for this purpose, and the method adopted was to lower into the water by a long tube a vertical circular disc of metal, at the centre of which was a box carrying a microphone. With such an arrangement the response is greatest when the sound waves strike the disc normally. Hence, if the disc is rotated about the tube as axis, there will be a certain plane of the disc for which the effect on the

microphone will be greatest. (The effect was indicated electrically by wires from the microphone passing up through the tube to an apparatus on the observing ship.) The direction of the submarine is thus revealed. To distinguish between the two opposite directions which were both normal to the disc, use was made of the fact that if a block of some non-resounding material of certain dimensions, containing an air space, is fixed to one side, sounds approaching that side have little, if any, effect on the microphone. There was thus only one position instead of two in the complete revolution of the disc through 360° , in which the sound swelled out.

Absorption of Sound

It is often more important to suppress sound than to amplify it; for example, in designing halls and buildings so that reverberation shall not interfere with good hearing. The problem here is to convert the energy of the molecular vibrations into some other form—usually the random motion of heat—so that it ceases to be propagated as sound. The method adopted for doing this is to interpose in the path of the sound some material producing high viscosity, so that the vibrations of each layer of molecules are checked by those of the adjacent layer, and are soon dissipated as heat in the medium. A porous material—such as cork or asbestos—is best for this purpose, for it consists of frequent alternations of gas and solid. We have already seen that when a fluid moves along a tube, the layer next to the wall remains at rest, so that the viscous drag there must be very great. The effect of a large number of solid/gas surfaces is therefore to damp out the vibrations rapidly. This is the principle on which sound-proof rooms are constructed, the surfaces of the rooms being covered with a porous material found by experiment to give satisfactory results.

Measurement of Sound

The complete measurement of sound would include the measurement of all its characteristics—pitch, quality, and

loudness. The first of these can be accurately measured by the frequency of the corresponding vibrations ; this subject has already been dealt with. The quality of sound does not lend itself to direct measurement. It is represented physically by the form of the wave-train, and this is best indicated either graphically or mathematically by the Fourier components into which the wave can be resolved. A single measure-number has no significance with such a characteristic.

The measurement of loudness, or intensity, is obviously important for practical purposes, but is difficult, if not impossible, to achieve satisfactorily. Since intensity corresponds to the amplitude of the sound wave, it would seem that it could be measured directly by the amplitude, but the matter is not so simple. In the first place, waves of the same amplitude but different frequency certainly do not give the same intensity to the ear ; in one case the noise may be objectionably loud, and in the other (in the supersonic region, for example) non-existent. Again, with waves of certain quality (see Fig. 53, page 221, for instance), it is not easy to say what the amplitude is.

No generally satisfactory method of measuring loudness, even for rough practical purposes, has yet been devised. What is important is usually not the physical characteristics of the sound wave, but the effect on the ear of the observer, and different observers vary widely in their appreciation of sound. They are more constant, however, in their judgment of the variation from one intensity to another. Thus, within the range of audible frequencies, the minimum *increase* of intensity which an observer can notice is a certain fraction of (not addition to) the original intensity, which is the same whatever the original intensity may be. It is thus possible to arrange a scale of intensity *ratios*. Suppose, as a simple example, that in two sound waves having the same pitch and quality but different amplitudes, the sound energies are respectively E_1 and E_2 . Then these energies can be compared by comparison of amplitudes (the energy in a sound

wave is, in fact, proportional to the square of the amplitude), and the ratio $\frac{E_1}{E_2}$ can be evaluated. The common logarithm of this ratio, $\log_{10} \frac{E_1}{E_2}$, is then said to express the energy ratio in *bels*, and gives a fairly objective indication of loudness difference. A more convenient expression is one-tenth of this—a quantity known as the *decibel*. The energy relation, L , of the two sounds is then defined by

$$L = 10 \log_{10} \frac{E_1}{E_2} \text{ decibels} \quad \dots \quad (17.1)$$

from which it follows that one decibel corresponds to an energy ratio, $\frac{E_1}{E_2}$, of 1.26. With a scale of this type of course, two decibels will not correspond to twice this energy ratio.

The decibel represents approximately the smallest increment of sound energy that the ordinary ear can detect. It must be clearly understood that it is a measure not of the actual intensity of a particular sound but of the ratio of intensity of two sounds.

Reproduction of Sound

There are two separate processes which may be described as the reproduction of sound. In the first—exemplified by the telephone and radio—the sound, as it is produced, is transformed into another form of energy—electrical—and transmitted as such over a distance, after which it is reconverted into sound energy again and the original tones are reproduced. This is a large subject, requiring separate treatment.

The second process is that by which the sound is made to create a wave-form in a piece of matter, in such a way that the original sound can be re-created again whenever required. In this process the original sound-energy is lost, but the record enables fresh energy, supplied from another source, to take the form of the sound first emitted. The

gramophone is the best-known instrument for this purpose. The sounds to be recorded are made to fall on a sensitive membrane carrying a style which rests on a uniformly moving drum or disc made of soft material, such as wax, on which the style can make an impression—a groove—whose depth depends on the intensity and whose form depends on the frequency and quality of the incident sound. When, later, a needle is made to travel over this impression, it is given up-and-down movements corresponding to the form of the groove, and if these movements are transmitted to a membrane, and thence to the air, the original sound waves are reproduced. They have, of course, to be amplified, and many refinements are introduced in order to obtain the best results, but the principle of the instrument is very simple.

EXERCISES

1. What is a *decibel*? Find the energy ratio, in decibels, of two sounds of the same pitch and quality, of which the amplitude of one is twice that of the other.
2. A constant source of sound is placed at a point A. Explain what you would do to ensure that its audibility at another point, B, is (a) as great, (b) as small as possible.

INDEX

Absolute temperature, 126, 157.
— zero, 127, 158.
Absorbing power, 185.
Absorption of sound, 240.
Acceleration, 13.
Acoustics of buildings, 232, 240.
Adiabatic changes, 150, 152, 212.
Adsorption, 78.
Alloys, 77, 109.
Altitude, 31.
— of pole, 32.
Amplification of sound, 238.
Amplitude (co-ordinate), 32.
— (of waves), 197, 210, 241.
Angle of contact, 67.
Anti-nodes, 203.
Apparent expansion, 109.
— solar time, 36.
Atmospheric pressure, 119.
Atoms, 18, 49.
Azimuth, 31.

Barometer, 119.
Beats, 219.
Bending of beams, 56.
Boiling point, 166, 174.
— — and pressure, 167, 173.
— — — solution, 173.
Bolometer, 186.
Boyle's law, 120.
Breaking-point, 53.
British system of units, 17, 139.
Bubbles, 71, 174.
Bulk modulus, 51.
Bunsen's ice calorimeter, 137.

Caloric, 97.
Calorie, 130.
Calorimeter, 134.
Capillarity, 70, 171.
Carbon dioxide, critical state of, 176.
Carnot cycle, 155.
Celestial equator, 22.
— latitude, 30.
— longitude, 30.
— poles, 22.

Celestial sphere, 21.
Centigrade temperature scale, 101.
Centimetre, 12.
C.G.S. system of units, 17.
Charles's law, 122.
Circumpolar stars, 24.
Clock, 37.
Coefficient of expansion, 106.
— — pressure increase, 122.
Cohesion, 63.
Colloids, 81.
Composition of vibrations, 220.
Compressibility of liquids, 60.
Compression, uniform, 50.
Concord, 219.
Conduction of heat, 102, 179.
Conductivity, thermal, 179.
Convection of heat, 103, 182.
Convective equilibrium, 182.
Co-ordinates, celestial, 27.
Critical pressure, 176.
— temperature, 174.
— velocity, 86.
— volume, 176.
Crova's disc, 194.
Crystalloids, 81.
Cubical expansion, 107.

Dalton's laws, 169.
Date-line, 40.
Decibel, 242.
Declination, 28.
Degree of temperature, 100.
Density, 110.
Detection of sound, 238.
Dewar flask, 186.
Dew point, 170.
Difference of principal specific heats, 149.
Diffusion, 74, 82.
Diffusivity, 83.
— coefficient, 85.
— (thermal), 179.
Dilatation, uniform, 50.
Discord, 219.
Doppler effect, 228.
Dynamics, 18.
Dyne, 14.

Ear, 238.
 Ear-trumpet, 232.
 East point, 31.
 Echoes, 231.
 Ecliptic, 26, 29.
 Efficiency of engine, 156.
 Effusion, thermal, 85.
 Elasticity, 50, 61.
 — perfect, 77.
 Elastic limit, 53.
 Emissive power, 185.
 End correction, 233.
 Energy, 15, 98.
 — conservation of, 16, 131.
 Entropy, 159.
 Equation of time, 37.
 — to wave-train, 197.
 Equinoxes, 43.
 Erg, 15.
 Ether, the, 103, 183.
 Evaporation, 84.
 Expansion, thermal, 105.
 External work, 142, 152.
 Fahrenheit temperature scale, 101.
 Fick's law, 83.
 First point of Aries, 26, 28.
 — — — Libra, 26.
 Fixed points of thermometer, 101.
 Flame, manometric, 239.
 Foot-pound, 17.
 Foot-poundal, 17.
 Force, 13.
 Forced vibrations, 206.
 Fourier's theorem, 200, 211, 222.
 Freezing mixture, 164.
 Freezing point, 164.
 Frequency, 195, 210.
 — measurement of, 217.
 — natural, 206.
 Friction, 74, 97.
 Fundamental note, 222, 234.
 g, 16, 90, 120.
 Gas constant, 126, 149.
 Gases, 49, 61, 117, 169.
 Gay Lussac's law, 122.
 Glacier motion, 165.
 Gram, 14.
 Gramophone, 211, 243.
 Gravitation, 15, 19, 65, 74, 90.
 — constant of, 15.
 Harmonic analysis, 201.
 Harmonics, 222, 227.
 Heat, 19, 97.
 Heat capacity, 134.
 — measurement of, 129.
 Hooke's law, 51, 53, 192, 201, 210.
 Horizon, 23, 30.
 Horse-power, 17.
 Hour, 28, 37.
 — angle, 32.
 — circles, 28.
 Humidity, 170.
 Hygrometers, 170.
 Icing, 165.
 Ideal gas, 123.
 Impact, 74, 75.
 Indicator diagram, 152.
 Inelasticity, perfect, 77.
 Intensity of sound, 210, 241.
 Interdiffusivity, 84.
 Interference of waves, 205, 234.
 Internal work, 142, 177.
 Irreversible change, 154.
 Isentropics, 160.
 Isometric change, 148.
 Isopiestic change, 148.
 Isothermal change, 149, 152, 175,
 212.
 Joule, 17.
 Kelvin temperature scale, 157.
 Kinetic energy, 15.
 — theory, 49, 76.
 Kundt's dust tube, 235.
 Latent heat, 130, 136, 172.
 Leslie's cube, 185.
 Linear expansion, 106.
 Liquefaction of gases, 176.
 Liquids, 49, 52, 60, 61, 87, 109.
 Local time, 39.
 Location of sound, 215.
 — — — submarines, 239.
 Longitudinal waves, 192, 210.
 Loops, 203, 234.
 Manometer, 119.
 Mass, 13.
 Matter, 49.
 Mean Solar time, 36.

Mean Sun, 36.
 Measurement, 11.
 Mechanical equivalent of heat, 131, 138, 140.
 — physics, 18.
 Mechanics, 18.
 Melde's experiment, 225.
 Melting, 130.
 — point, 164.
 — — and pressure, 165.
 Meridian, 27.
 Meridians of longitude, 28.
 Method of mixtures, 134.
 Metre, standard, 12.
 Microphone, 238, 239.
 Midnight sun, 44.
 Minute, 29.
 Mixture, 78.
 Moduli of elasticity, 51, 195, 204, 215.
 Molecular theory, 49, 61, 64, 82, 84, 88, 91, 99, 123, 162, 182.
 Molecules, 18, 49.
 Moment of inertia of cross-section, 58.
 Momentum, 14, 76.
 Monochord, 224.
 Motion, 11, 13, 98.
 — laws of, 13, 14, 16.
 Movement of stars, 21.
 Nadir, 31.
 Nodes, 203, 234.
 North point, 31.
 N.T.P., 120, 213.
 Nuclci, condensation on, 172, 174.
 — solidification on, 165.
 Occlusion, 78.
 Optic axis, 223.
 Osmosis, 74, 80.
 Osmotic pressure, 80, 81.
 Overtones, 222, 227, 234.
 Parallels of declination, 28.
 — — latitude, 28.
 Partial pressure, 169.
 Perfect gas, 123.
 — — equation, 125.
 Period of wave, 198.
 Phase, 199.
 — change on reflection, 203, 234.
 Physics, 11.
 Pipes, sound in, 215, 232.
 Pitch, 210.
 Poiseuille's formula, 88, 91.
 Poisson's ratio, 52.
 Poles, celestial, 22.
 — of ecliptic, 29.
 Porous plug experiment, 142.
 Potential energy, 65, 131.
 Power, 17.
 Pressure, 50, 91, 112, 117, 166.
 — in bubbles, 71.
 Prévost's theory of exchanges, 184.
 Prime vertical, 31.
 Principal specific heats, 139.
 Production of sound, 237.
 Properties of matter, 19, 47.
 Quality of sound, 210.
 Quartz oscillator, 223, 238.
 Radiant energy, 74, 183.
 Radiation, 102, 183.
 Radio, 242.
 Range of audibility, 222.
 Ratio of principal specific heats, 141, 151, 213.
 Real expansion, 109.
 Réaumur temperature scale, 101.
 Reflecting power, 186.
 Reflection of sound, 231.
 — — waves, 201, 205.
 Reproduction of sound, 242.
 Resonance, 207, 232.
 Resonators, 227, 238.
 Restitution, coefficient of, 76.
 Reversible change, 154.
 Right ascension, 28.
 Rigidity, 54, 87.
 — modulus, 51.
 Saturated solutions, 79.
 — vapours, 169.
 Saturation vapour pressure, 166.
 Seasons, 41.
 Second, 29.
 — mean solar, 12.
 Semi-permeable membranes, 80.
 Shear, 50, 87.
 — angle of, 51.
 Sidereal time, 25, 33.
 Simple harmonic motion, 201.
 Siren, 237.

Solar time, 35.
 Solids, 49, 52, 61, 105.
 Solstices, 43.
 Solute, 79.
 Solution, 74, 77, 164.
 Solvent, 79.
 Sound, 191, 209.
 — measurement of, 240.
 South point, 31.
 Space, 12.
 Specific heat, 133.
 — — at constant pressure, 140.
 — — — volume, 141.
 Standard meridian, 39.
 — time, 39.
 Statics, 18.
 Stationary waves, 202, 234.
 Steady motion, 86.
 Stefan-Boltzmann law, 186.
 Stokes's law, 89.
 Strain, 50.
 Stress, 50.
 Stroboscope, 217.
 Sublimation, 166, 170.
 Summer time, 40.
 Sundial, 38.
 Sun, motion of, 25.
 Supercooling, 165.
 Supersaturated solutions, 79.
 Supersonics, 223, 238.
 Surface energy, 65, 67.
 — radiation, 184.
 — tension, 63.
 Telephone, 242.
 Temperature, 92, 98.
 — measurement of, 99, 157.
 Tensile strength of liquids, 60.
 Tension, 50.
 Thermodynamics, first law of, 132, 145.
 — second law of, 161.
 Thermodynamic temperature scale, 157.
 Thermometer, gas, 101, 127.
 — mercury-in-glass, 100.
 Thermopile, 186.
 Time, 12.
 Total heat, 172.
 Transfer of heat, 102, 129, 161, 179.
 Transformation of energy, 132, 145, 155, 242.
 Transit circle, 35.
 Transverse waves, 192, 196, 210.
 Tuning fork, 218, 226, 237.
 Turbulent motion, 86.
 Ultracentrifuge, 90.
 Units, 17.
 Unsaturated vapours, 169.
 Valve, thermionic, 238.
 Vapour pressure, 166, 168.
 — — and surface curvature, 171.
 — — of ice, 170.
 Vapours, 169.
 Velocity, 13.
 — of sound, 151, 212, 236.
 — — waves, 194.
 Vertical circle, 31.
 Vibrating strings, 224.
 Vibrations, 191, 206, 209.
 Violin, 225.
 Viscosity, 74, 85, 240.
 — coefficient of, 87.
 Water, expansion of, 116.
 — value of calorimeter, 136.
 Watt, 17.
 Wave-length, 194.
 — motion, 191.
 — train, 193.
 Weight, 15.
 West point, 31.
 Whispering galleries, 231.
 Work, 16, 131.
 Working substance of engine, 157.
 Yard, standard, 12.
 Young's modulus, 51, 54.
 Zenith, 27, 31.
 Zodiac, 26.

